OLUME 107 PART C NUMBER 11 MARCH 1960

J. T. LUDWIG
MAY 3 1960
SCIENTIFIC ADVISORY GROUP

The Proceedings

OF

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871: INCORPORATED BY ROYAL CHARTER 1921

PART C

MONOGRAPHS Nos. 340–353

SAVOY PLACE · LONDON W.C.2

Price Fifteen Shillings

The Institution of Electrical Engineers

FOUNDED 1871
INCORPORATED BY ROYAL CHARTER 1921

PATRON: HER MAJESTY THE QUEEN

COUNCIL 1959–1960

President

SIR WILLIS JACKSON, D.Sc., F.R.S.

Past-Presidents

W. H. ECCLES, D.Sc., F.R.S.
THE RT. HON. THE EARL OF MOUNT EDGCUMBE, T.D.
J. M. DONALDSON, M.C.
PROF. E. W. MARCHANT, D.Sc.
H. T. YOUNG.
SIR GEORGE LEE, O.B.E., M.C.
SIR ARTHUR P. M. FLEMING, C.B.E., D.Eng., LL.D.
J. R. BEARD, C.B.E., M.Sc.
SIR NOEL ASHBRIDGE, B.Sc.(Eng.).
SIR HARRY RAILING, D.Eng.
P. DUNSHEATH, C.B.E., M.A., D.Sc.(Eng.), LL.D.
SIR VINCENT Z. DE FERRANTI, M.C.
T. G. N. HALDANE, M.A.
PROF. E. B. MOULLIN, M.A., Sc.D., LL.D.
SIR ARCHIBALD J. GILL, B.Sc.(Eng.).
SIR JOHN HACKING.
COL. B. H. LEESON, C.B.E., T.D.
SIR HAROLD BISHOP, C.B.E., B.Sc.(Eng.), F.C.G.I.
SIR JOSIAH ECCLES, C.B.E., D.Sc.
THE RT. HON. THE LORD NELSON OF STAFFORD.
SIR GORDON RADLEY, K.C.B., C.B.E., Ph.D.(Eng.).
S. E. GOODALL, M.Sc.(Eng.), F.Q.M.C.

Vice-Presidents

O. W. HUMPHREYS, C.B.E., B.Sc.
G. S. C. LUCAS, O.B.E., F.C.G.I.
SIR HAMISH D. MACLAREN, K.B.E., C.B., D.F.C., LL.D., B.Sc.
C. T. MELLING, C.B.E., M.Sc.Tech.
A. H. MUMFORD, O.B.E., B.Sc.(Eng.).

Honorary Treasurer

E. LEETE.

Ordinary Members of Council

PROF. H. E. M. BARLOW, Ph.D., B.Sc.(Eng.).
C. O. BOYSE, B.Sc.(Eng.).
PROF. M. W. HUMPHREY DAVIES, M.Sc.
SIR JOHN DEAN, B.Sc.
L. DRUCQUER.
J. M. FERGUSON, B.Sc.(Eng.).
D. C. FLACK, B.Sc.(Eng.), Ph.D.
J. S. FORREST, D.Sc., M.A.
R. J. HALSEY, C.M.G., B.Sc.(Eng.), F.C.G.I.
J. B. HIGHAM, Ph.D., B.Sc.
R. A. HORE, M.A., B.Sc.
F. C. MCLEAN, C.B.E., B.Sc.
B. L. METCALF, B.Sc.(Eng.).
J. R. MORTLOCK, Ph.D., B.Sc.(Eng.).
THE HON. H. G. NELSON, M.A.
R. H. PHILLIPS, T.D.
H. V. PUGH.
J. R. RYLANDS, M.Sc., J.P.
G. A. V. SOWTER, Ph.D., B.Sc.(Eng.).
C. E. STRONG, O.B.E., B.A., B.A.I.
D. H. TOMPSETT, B.Sc.(Eng.).

Chairmen and Past-Chairmen of Sections

Electronics and Communications:
M. J. L. PULLING, C.B.E., M.A.
*G. MILLINGTON, M.A., B.Sc.

Measurement and Control:
PROF. A. TUSTIN, M.Sc.
*J. K. WEBB, M.Sc.(Eng.), B.Sc.Tech.

Supply:
J. R. MORTLOCK, Ph.D., B.Sc.(Eng.).
*D. P. SAYERS, B.Sc.

Utilization:
T. E. HOUGHTON, M.Eng.
*R. A. MARRYAT, B.Sc.(Eng.).

Chairmen and Past-Chairmen of Local Centres

East Midland Centre:
D. H. PARRY, B.Sc.
*D. E. LAMBERT, B.Sc.(Eng.).

Mersey and North Wales Centre:
*T. A. P. COLLEDGE, B.Sc.(Eng.).
J. COLLINS.

North Midland Centre:
PROF. G. W. CARTER, M.A.
*J. D. NICHOLSON, B.Sc.

North-Eastern Centre:
H. WATSON-JONES, M.Eng.
*A. T. CRAWFORD, B.Sc.

North-Western Centre:
F. J. HUTCHINSON, M.Eng.
*PROF. F. C. WILLIAMS, O.B.E., D.Sc., D.Phil., F.R.S.

Northern Ireland Centre:
T. S. WYLIE.
*D. S. MCILLHAGGER, Ph.D., M.Sc.

Western Centre:
H. JACKSON, B.Sc.(Eng.).
*R. W. STEEL.

Scottish Centre:
J. A. AKED, M.B.E.
*R. J. RENNIE, B.Sc.

South Midland Centre:
G. F. PEIRSON.
*L. L. TOLLEY, B.Sc.(Eng.).

Southern Centre:
W. D. MALLINSON, B.Sc.(Eng.).
*G. BISHOP, B.Sc.

* Past Chairman.

Secretary

W. K. BRASHER, C.B.E., M.A., M.I.E.E.

Deputy Secretary

F. JERVIS SMITH, M.I.E.E.

Principal Assistant Secretary

F. C. HARRIS.

Editor-in-Chief

G. E. WILLIAMS, B.Sc.(Eng.), M.I.E.E.

The Institution is not, as a body, responsible for the opinions expressed by individual authors or speakers.
An example of the preferred form of bibliographical references will be found beneath the list of contents.

THE PROCEEDINGS OF
THE INSTITUTION OF ELECTRICAL ENGINEERS

EDITED UNDER THE SUPERINTENDENCE OF W. K. BRASHER, C.B.E., M.A., M.I.E.E., SECRETARY

VOL. 107. PART C. No. 11. MARCH 1960

DISCUSSION ON 'EDDY-CURRENT LOSSES IN THIN FERROMAGNETIC SHEETS'*

Mr. A. C. Sim (*communicated*): The paper gives an expression, ue to Polivanov, for the ratio of actual eddy-current loss to at expected by the usual formulae. This is the series

$$\eta = \frac{96\phi}{\pi^3} \sum_{n=1}^{\infty} \frac{\coth[(2n-1)\pi\phi]}{(2n-1)^3} \quad . \quad . \quad (A)$$

ne author has plotted this and produced an empirical expression

$$\eta \simeq (1 + kd^{-r}) \quad . \quad . \quad . \quad . \quad . \quad (B)$$

here r lies between 1·5 and 1·7, which is claimed to be fairly curate for ϕ less than unity, where ϕ represents the ratio, d, of domain size to sheet thickness.

No simple expression is given for the series (A) that is more nvenient when ϕ is small, so this is offered here.

Consider the integral

$$\int \frac{\coth(2\pi z\phi)\tan(\pi z)}{z^3} dz \quad . \quad . \quad . \quad . \quad (C)$$

ken over an infinite circle in the complex z-plane, chosen so to avoid any poles. Since the integrand vanishes at every oint on the circle the integral is zero, and hence the sum of sidues may be equated to zero.

The residue at the pole $z = 0$ is

$$2\pi j\pi^2(1 + 4\phi^2)/6\phi$$

at at the pole $z = \frac{1}{2}(2m - 1)$ is

$$-2\pi j \frac{8\coth[(2m-1)\pi\phi]}{\pi(2m-1)^3}$$

d that at the pole $z = jn/2\phi$ is

$$-2\pi j \frac{4\phi^2 \tanh(n\pi/2\phi)}{\pi n^3}$$

Summing these residues for $m = 0, \pm 1, \pm 2, \ldots$, and $= \pm 1, \pm 2, \ldots$, and equating the result to zero gives the ernative expansion

$$\eta = \left[1 + 4\phi^2 - \frac{48\phi^3}{\pi^3} \sum_{n=1}^{\infty} \frac{\tanh(n\pi/2\phi)}{n^3}\right] \quad . \quad . \quad (D)$$

When ϕ is less than unity this approximates to terms in ϕ^2 and ϕ^3 only. By replacing the hyperbolic tangent by an exponential series and changing the order of summation, a further form is found in terms of trilogarithms:*

$$\eta = [1 + 4\phi^2 - 48\zeta(3)\phi^3/\pi^3 + (96\phi^3/\pi^3) \sum_{n=1}^{\infty} (-1)^{n+1} Li_3(\varepsilon^{-n\pi/\phi})] \quad (E)$$

For ϕ less than unity, negligible error arises from

$$(\eta - 1) \simeq 4\phi^2\{1 - (12\phi/\pi^3)[\zeta(3) - 2\varepsilon^{-\pi/\phi}]\} \quad (F)$$

where $12/\pi^3 = 0{\cdot}38\,702$ and $\zeta(3) = 1{\cdot}202$.

I will not take up space with a graph, but I do not consider that anyone plotting formula (F) on double logarithmic paper would agree with the range 1·5 to 1·7 for r in the approximation (B). The most appropriate range appears to be 1·7–1·8 for ϕ between 0·1 and 1·0.

This is important, of course, should anyone wish to compare with experiment by plotting $(\eta - 1)$ against d. If a slope of 1·5 were found, this, I feel, would indicate a significant departure from Polivanov's theory, and could be used to imply a probable law of dependence of domain size upon sheet thickness, or else a departure from Polivanov's model.

Mr. E. W. Lee (*in reply*): Mr. Sim is, of course, quite right; the exponent r given in the paper is incorrect and may be misleading. My point was that Polivanov's theory, which gives a smooth curve for η as a function of ϕ should not be considered inadequate merely because the experimental data happen to fit a simple power law, as this itself could easily be a consequence of the restricted range of ϕ over which the measurements were made.

Since the paper was written this subject has received attention from numerous authors. My present feeling is that the essential correctness of Polivanov's theory is now well established, and that a comparison of the theory with experiment would enable the domain size to be estimated within a factor of about two. By extending the frequency range of the measurements it might even be possible to ascertain the distribution of domain sizes. Such considerations, however, fall outside the intended scope of the paper.

* LEE, E. W.: Monograph No. 284 M, February, 1958 (see 105 C, p. 337).

* LEWIN, L.: 'Dilogarithms and Associated Functions' (Macdonald, London, 1958), p. 258.

1960: The Institution of Electrical Engineers

621.313.32

The Institution of Electrical Engineer
Monograph No. 340
Aug. 195

A FREQUENCY-RESPONSE METHOD FOR THE PREDETERMINATION OF SYNCHRONOUS-MACHINE STABILITY

By A. S. ALDRED, M.Sc., Associate Member, and G. SHACKSHAFT, B.Eng., Ph.D., Graduate.

(*The paper was first received* 21*st November*, 1958, *and in revised form* 23*rd March*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in August*, 1959.)

SUMMARY

The paper presents a new concept for the predetermination of synchronous-machine stability. The concept is based on the realization of a basic closed-loop pattern for a synchronous generator, which, when established, can be subjected to the frequency-response procedure of the Nyquist stability criterion. The basic closed-loop pattern emerges from the application of small-displacement theory to Park's equations for a synchronous generator. The method is applicable to a machine with or without a voltage regulator, but is more useful in the latter case, and this forms the majority of the applications presented in the paper.

It is shown that, as a by-product of the analysis, expressions for the damping in a synchronous machine may be derived. Since this is considered to be of some interest, the influence of some of the more important parameters of the machine and regulator on the damping coefficient is demonstrated.

LIST OF SYMBOLS

δ = Rotor angle.
f, ω = Frequency, c/s and rad/s.
H = Inertia constant, kWs/kVA.
$M = H/180f$.
T_i, P_i = Torque and power input.
T_u, P_u = Torque and power output.
$p\theta$ = Speed.
V_{fd} = Generator field voltage.
R_{fd} = Generator field resistances.
I_{fd} = Generator field current.
X_{fd} = Generator field reactance.
X_{ad} = Mutual reactance between generator field and direct-axis armature winding.
$V_f = \frac{V_{fd}}{R_{fd}} X_{ad} p\theta$, open-circuit excitation voltage.
Φ_{fd} = Field flux linkage.
Φ_{dm} = Machine direct-axis flux linkage.
Φ_{qm} = Machine quadrature-axis flux linkage.
$\tau_{d0} = X_{fd}/R_{fd}$ = Generator field time-constant.
K_d = Damping torque coefficient.
V_b = Busbar voltage.
V_{dm}, V_d = Direct-axis machine terminal voltage and busbar voltage, respectively.
V_{qm}, V_q = Quadrature-axis machine terminal voltage and busbar voltage, respectively.
I_d, I_q = Direct- and quadrature-axis currents.
X_{dm}, X_{qm} = Machine direct- and quadrature-axis synchronous reactances.
X_t = Transmission-line series reactance.
X_d, X_q = Total direct- and quadrature-axis reactances.
X'_{dm} = Machine direct-axis transient reactance.
X'_d = Total direct-axis transient reactance.
V_m = Machine terminal voltage.
V_r = Reference voltage.
μ = Regulator loop gain.
τ_s, μ_s = Stabilizer time-constant and amplification factor.
τ_e = Exciter field time-constant.
K = Amplification factor in the positive-feedback loop

Correspondence on Monographs is invited for consideration with a view to publication.

Mr. Aldred and Dr. Shackshaft are in the Department of Electrical Engineering, University of Liverpool.

(1) INTRODUCTION

Steady-state stability is normally deduced from consideratio of the power-angle curve. This procedure ignores dampin both positive and negative, in the machine and consequenti does not provide a complete picture or solution of the stabili problem. This is particularly true for a machine with a volta regulator, when operating in the dynamic-stability region wi positive synchronizing torque, because the movement of t rotor, subsequent to a small displacement, is largely depende upon the damping in the system. If positive damping present the rotor oscillations will die away. If there is no dam ing the machine will hunt, and if the damping is negative t rotor oscillations will increase in amplitude and may eventual cause the machine to become unstable. A more exact metho which has been employed by Concordia,[1] is to determine t stability by the application of Routh's[2] criterion to the coefficien of the characteristic equation of motion of the machine. Becau the synchronous-machine characteristic equation is non-linear is first necessary to linearize it by the method of small displac ments, initially described by Park[3] in relation to synchrono machines. The Routh stability criterion, however, suffers fro the general disadvantage that it provides no information concer ing the degree of stability of the machine; this knowledge frequently more important than merely knowing whether, und certain operating conditions, the machine is stable or not. was observed by the authors that the logical extension of Park small-oscillation theory led naturally to the application of t Nyquist[4] criterion, with its attendant advantages of providi information concerning degree of stability and a rapid appreci tion of the effects of modifications of the machine equatio owing to the addition of, for example, damper windings or voltage regulator. The Nyquist criterion is a universally-appli closed-loop control-system technique by which stability deduced from a graphical plot of the open-loop frequen response. This implies the physical existence of a main lo that can be broken at some convenient point in order to det mine the frequency response, either analytically or expe mentally. When the synchronous machine is considered fr this aspect the question arises as to what is the principal lo and what constitutes the feedback. The answers are fou from an examination of the small-displacement equation motion, which is in terms of either torque or power. The inp to the closed loop is the torque from the prime mover ΔT_i, a the output is the rotor angle $\Delta\delta$. The latter is operated on a function which may be written as $f(p)$ in operational form $f(j\omega)$ in frequency form, to provide the feedback $f(j\omega)\Delta\delta$, wh is, in fact, the sum of the damping and synchronizing torqu

may be shown that the addition of a voltage regulator, damper ndings or any change in the electrical parts of the machine, ily alters the form of the function $f(j\omega)$ and does not interfere th the basic closed-loop pattern. The essential difference, so r as stability is concerned, between the synchronous machine d a conventional closed-loop control system, is that, in the rmer, the main loop cannot be broken, i.e. the feedback is herent; consequently the Nyquist criterion can only be applied alytically and not experimentally, unless recourse is had to a nulator of appropriate form.[5,6,7]

The objects of the paper are to develop the frequency-response ethod and use it to demonstrate the influence of some of the ore important factors on the stability of a machine without, it primarily with, a voltage regulator. An examination will so be made of the damping in a system, and it will be shown at damping torques can be evaluated as a by-product of the equency-response analysis. In the authors' analysis the sources damping are in the field and the voltage regulator. There are her sources of damping that may be important, e.g. positive mping arising from the use of a damper winding and negative mping associated with armature resistance. The authors' ethod may be extended to include, if necessary, these additional urces.

2) SMALL-OSCILLATION THEORY FOR SYNCHRONOUS MACHINE WITH VOLTAGE REGULATOR

The system to be analysed is shown in block-schematic form Fig. 1. The operational equations for this system are derived Section 8 and are as follows:

$$V_d = X_q I_q \quad (1)$$
$$V_q = G(p)V_f - X_d(p)I_d \quad (2)$$
$$V_d = V_b \sin\delta \quad (3)$$
$$V_q = V_b \cos\delta \quad (4)$$
$$T_i = Mp^2\delta + K_d p\delta + \Phi_d I_q - \Phi_q I_d \quad (5)$$
$$V_f = g(p)(V_m - V_r) \quad (6)$$
$$V_m^2 = V_{dm}^2 + V_{qm}^2 \quad (7)$$
$$V_{dm} = V_d - X_t I_q \quad (8)$$
$$V_{qm} = V_q + X_t I_d \quad (9)$$

he frequency-response method is based on the small-oscillation uations for a synchronous machine. These are obtained by lowing all the variables to change by small amounts from e initial operating condition. Steady-state initial-condition ltages, currents and fluxes are denoted by the suffix 0.

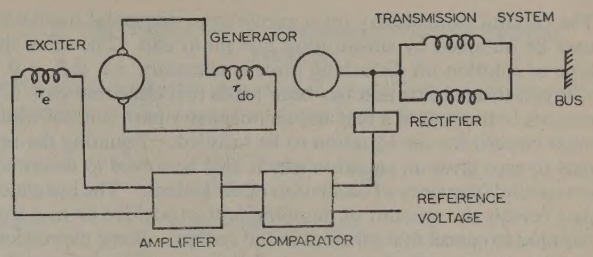


Fig. 1.—Basic system of synchronous machine and voltage regulator.

From eqn. (1),
$$\Delta V_d = X_q \Delta I_q \quad (10)$$
From eqn. (2),
$$\Delta V_q = G(p)\Delta V_f - X_d(p)\Delta I_d \quad (11)$$
From eqn. (3),
$$\Delta V_d = V_b \cos\delta_0 \Delta\delta \quad (12)$$
From eqn. (4),
$$\Delta V_q = -V_b \sin\delta_0 \Delta\delta \quad (13)$$
From eqn. (5),
$$\Delta T_i = \Phi_{d0}\Delta I_q + I_{q0}\Delta\Phi_d - \Phi_{q0}\Delta I_d - I_{d0}\Delta\Phi_q + Mp^2\Delta\delta + K_d p\Delta\delta \quad (14)$$
Now
$$\Delta V_d = -\Delta\Phi_q \quad (15)$$
and
$$\Delta V_q = \Delta\Phi_d \quad (16)$$
Therefore
$$\Delta T_i = I_{q0}G(p)\Delta V_f - [I_{q0}X_d(p) + \Phi_{q0}]\Delta I_d + [\Phi_{d0} + I_{d0}X_q]\Delta I_q + Mp^2\Delta\delta + K_d p\Delta\delta \quad (17)$$
From eqn. (6),
$$\Delta V_f = g(p)\Delta V_m \quad (18)$$
Substituting for ΔV_m by using eqns. (7), (8) and (9) we get
$$\Delta V_f = g(p)[V_d'\Delta V_d - V_d' X_t \Delta I_q + V_q'\Delta V_q + V_q' X_t \Delta I_d] \quad (19)$$
where
$$V_d' = \frac{V_{dm0}}{V_{m0}} \quad (20)$$
and
$$V_q' = \frac{V_{qm0}}{V_{m0}} \quad (21)$$

If eqn. (19) is substituted in eqns. (11) and (14) the complete set of small-displacement equations for the machine and voltage regulator may be written in matrix form, as in eqn. (22).

$$\begin{vmatrix} \Delta V_d \\ \Delta V_q \\ \Delta I_d \\ \Delta I_q \\ \Delta\delta \end{vmatrix} \times \begin{vmatrix} I_{q0}G(p)g(p)V_d' & I_{q0}G(p)g(p)V_q' & G(p)g(p)I_{q0}X_tV_q' - \Phi_{q0} - X_d(p)I_{q0}' & \Phi_{d0} + I_{d0}X_q - G(p)g(p)I_{q0}X_tV_d' & Mp^2 + K_dp \\ -1 & 0 & 0 & X_q & 0 \\ G(p)g(p)V_d' & G(p)g(p)V_q' - 1 & G(p)g(p)V_q'X_t - X_d(p) & -G(p)g(p)V_d'X_t & 0 \\ -1 & 0 & 0 & 0 & V_b\cos\delta_0 \\ 0 & +1 & 0 & 0 & V_b\sin\delta_0 \end{vmatrix} = \begin{vmatrix} \Delta T_i \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix} \quad (22)$$

Then eqn. (22) is solved for the ratio $\Delta T_i/\Delta\delta$ it yields

$$\frac{\Delta T_i}{\Delta\delta} = Mp^2 + K_d p + \frac{V_b}{X_d}V_{f0}\cos\delta_0 + \frac{V_b^2(X_d - X_q)}{X_d X_q}\cos 2\delta_0 + \frac{+ G(p)g(p)V_b\sin\delta_0[X_{dm}X_qV_b\sin\delta_0 V_q' - X_{qm}X_dV_b\cos\delta_0 V_d'] + V_b^2\sin^2\delta_0 X_q[X_d(p) - X_d]}{X_dX_q[G(p)g(p)V_q'X_t - X_d(p)]} \quad (23)$$

The conditions necessary for a continuous sinusoidal oscillation may be obtained by substituting $p = j\omega$ in eqn. (23). For this type of solution no disturbing force is necessary, i.e. $\Delta T_i = 0$.

When this substitution has been made it is clear that eqn. (23) reduces to the sum of a real and an imaginary part, both of which must be zero for the equation to be satisfied. Equating the real part to zero gives an equation which may be solved to determine the natural frequency of oscillation of the system. The imaginary part reveals the amount of damping (either positive or negative) required to cancel that inherent in the system. Some expressions for this damping coefficient are as follows:

(*a*) *Machine without Regulator*

$$G(p) = \frac{1}{1 + \tau_{d0}p}, \quad X_d(p) = \frac{X_d + X'_d\tau_{d0}p}{1 + \tau_{d0}p}, \quad g(p) = 0$$

$$K_d = \frac{\tau_{d0}(X_d - X'_d)V_b^2 \sin^2 \delta_0}{X_d^2 + \omega^2\tau_{d0}^2 X_d'^2} \quad . \quad . \quad . \quad (24)$$

(*b*) *Machine with Regulator*

$$G(p) = \frac{1}{1 + \tau_{d0}p}, \quad X_d(p) = \frac{X_d + X'_d\tau_{d0}p}{1 + \tau_{d0}p}, \quad g(p) = \frac{-\mu}{1 + \tau_e p}$$

$$K_d = \frac{\dfrac{V_b^2 \sin \delta_0}{X_q}\begin{bmatrix}\mu\tau'_{d0}(V'_d X'_d X_{qm} \cos \delta_0 - V'_q X'_{dm} X_q \sin \delta_0 \\ - \mu\tau_e(V'_d X_d X_{qm} \cos \delta_0 - V'_q X_{dm} X_q \sin \delta_0) \\ + (X_d - X'_d)X_q \sin \delta_0 \tau_{d0}(1 + \omega^2\tau_e^2)\end{bmatrix}}{[(\mu V'_q X_t + X_d - X'_d\tau_e\tau_{d0}\omega^2)^2 + \omega^2(\tau_e X_d + \tau_{d0}X'_d)^2]}$$

. . . . (25)

(*c*) *Machine with Regulator incorporating a Positive Feedback proportional to Rate of Change of Field Current.*

$$G(p) = \frac{1}{1 + (\tau_{d0} - K)p}, \quad X_d(p) = \frac{X_d + (\tau_{d0}X'_d - KX_d)p}{1 + (\tau_{d0} - K)p},$$

$$g(p) = -\mu$$

in which K is a measure of the positive feedback.

$$K_d = \frac{\dfrac{V_b^2 \sin \delta_0}{X_q}\begin{bmatrix}\mu\tau_{d0}(V'_d X'_d X_{qm} \cos \delta_0 - V'_q X_q X'_{dm} \sin \delta_0) \\ - \mu K(V'_d X_{qm} X_d \cos \delta_0 - V'_q X_q X_{dm} \sin \delta_0) \\ + \tau_{d0}(X_d - X'_d)X_q \sin \delta_0\end{bmatrix}}{(\mu V'_q X_t + X_d)^2 + \omega^2(\tau_{d0}X'_d - KX_d)^2}$$

. . . . (26)

These expressions for the damping coefficient have been obtained as a by-product of the frequency-response method. The damping characteristics are considered to be of interest, but as this is only an intermediate state in the development of the frequency-response method a discussion of the damping is deferred until a later Section.

(3) THE APPLICATION OF NYQUIST'S CRITERION OF STABILITY TO THE SMALL-OSCILLATION EQUATIONS

(3.1) Nyquist Criterion

The Nyquist criterion of stability is used to predetermine the stability or otherwise of closed-loop systems of either single- or multi-loop configuration. For single-loop systems the application of the criterion is relatively straightforward. For multi-loop systems, however, this is not so, since a situation might arise in which one or more of the subsidiary loops in the system may cause the system to be unstable when the main loop is opened; the presence of the main loop, in fact, is necessary for the maintenance of stability of the complete system. To study the general case, consider the closed-loop system of Fig. 2, whic consists of a forward transfer function, $\phi_1(p)$, and a feedbac transfer function, $\phi_2(p)$. It is assumed that $\phi_1(p)\phi_2(p)$ ma

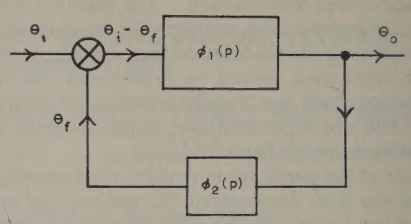


Fig. 2.—Fundamental closed-loop system.

represent either a single- or a multi-loop system. The relatio ship between the controlled quantity and the controllin quantity is

$$\frac{\theta_o}{\theta_i} = \frac{\phi_1(p)}{1 + \phi_1(p)\phi_2(p)} \quad . \quad . \quad . \quad . \quad (2$$

By inspection of eqn. (27) it is clear that, for stability, t function $[1 + \phi_1(p)\phi_2(p)]$ must have no zeros with positive re parts.

To develop the Nyquist criterion use is made of an extensio of a theorem in complex variable mathematics due to Cauch Cauchy's theorem states that, if a function f(*z*) is analytic c and within a closed contour on the *z*-plane,

$$\oint f(z) = 0$$

The extension of the theorem states that if a function is analyti except for possible poles, within and on a given contour, t number of times the plot of f(*z*) encircles the origin of the f(plane in an anti-clockwise direction, while *z* itself moves rou the prescribed contour once in a clockwise direction, is equal the number of poles of f(*z*) minus the number of zeros of f(within the contour. In stability studies the complex variable is, of course, the complex frequency p and the function f(*z*) is f(

The contour which is chosen in the p-plane when stability of interest in an infinite semicircle, as shown in Fig. 3. T

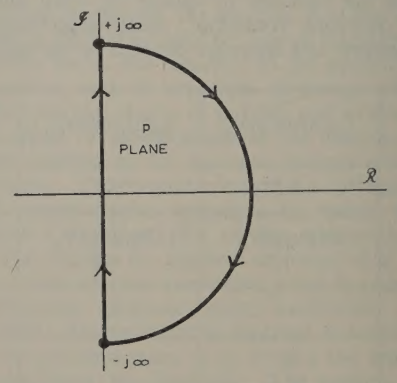


Fig. 3.—Infinite semicircle. Locus of p for Nyquist criterion.

reason for choosing this contour is that it encloses any poles the function $\phi_1(p)\phi_2(p)$ and consequently of the functi $1 + \phi_1(p)\phi_2(p)$ which have positive real parts, and also a zeros of $1 + \phi_1(p)\phi_2(p)$ with positive real parts. A pole of t function $\phi_1(p)\phi_2(p)$ with a positive real part implies that t

stem of Fig. 2 is unstable when the main loop is opened. les of the function $\phi_1(p)\phi_2(p)$ with negative real parts are not portant, since they represent stability of Fig. 2 when the main op is opened. The general case is considered in Fig. 3. The finite semicircle either encloses poles of $\phi_1(p)\phi_2(p)$ (representing unstable open loop) or it does not (representing a stable open op).

Since all physical systems have zero response at infinite frequency, the movement of the function $f(p)$ in the $f(p)$-plane will zero for infinite frequencies, both real and complex. The mplex frequencies represented by the semicircular path between $j\infty$ and $+j\infty$ are therefore unimportant, since, for these equencies, $f(p)$ is stationary. The frequencies which are important are the real frequencies between $-j\infty$ and $+j\infty$.

For practical purposes it is more convenient to consider the ot of $\phi_1(p)\phi_2(p)$ for $p = -j\infty$ to $+j\infty$ rather than $1 + (p)\phi_2(p)$, since the former is readily obtainable as the open-op transfer function. Consequently, if this function is plotted, e extension of Cauchy's theorem is modified to the statement at, for $1 + \phi_1(p)\phi_2(p)$ to have no zeros with positive real rts, the number of encirclements of the $(-1, 0)$ point by $(p)\phi_2(p)$ as p varies from $-j\infty$ to $+j\infty$ must equal the number poles of $\phi_1(p)\phi_2(p)$ with positive real parts.

If the function $\phi_1(p)\phi_2(p)$ is such that it is zero or finite when $= 0$ no difficulty is encountered in observing the number of t encirclements of the point $(-1, 0)$. In many cases, however, p tends to $j0$ the function $\phi_1(p)\phi_2(p)$ tends to c/p^n, where c a constant and n may have values of 1, 2 or 3 in practical cases. these cases as p tends to $j0$, $\phi_1(p)\phi_2(p)$ tends to infinity, and order to obtain values of $\phi_1(p)\phi_2(p)$ in this region it is necesry to plot this function as p moves round the origin in a very all semicircle in the positive half-plane as shown in Fig. 4.

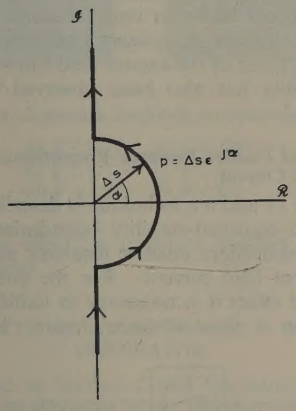


Fig. 4.—Locus of p near the origin for Nyquist criterion.

In general, as p tends to $j0$, substituting the value of p on the all semicircular path gives

$$\phi_1(p)\phi_2(p) = \frac{c}{\Delta s^n}\varepsilon^{-jn\alpha}$$

om this equation it may be seen that the magnitude of $c/\Delta s^n$ nds to infinity as Δs tends to zero, and the angle of $\phi_1(p)\phi_2(p)$ akes n clockwise semicircles centred on the origin as α passes rough the values $-\pi/2$ to $+\pi/2$. The Nyquist criterion has en dealt with at length, because, in its application to synronous machines, the closures are important, and under certain rcumstances there may be poles of the open-loop transfer nction with positive real parts.

(3.2) The Application to Synchronous-Machine Stability Problems

The small-oscillation equation of a synchronous machine and regulator [eqn. (23)] may be written in the form

$$\frac{\Delta T_i}{\Delta\delta} = Mp^2 + f(p) \quad . \quad . \quad . \quad . \quad . \quad (28)$$

where $$f(p) = \phi(\delta_0) + \phi(p) \quad . \quad . \quad . \quad . \quad . \quad (29)$$

From eqn. (23),

$$\phi(\delta_0) = \frac{V_b V_{f0}}{X_d}\cos\delta_0 + V_b^2\frac{(X_d - X_q)}{X_d X_q}\cos 2\delta_0 \quad . \quad (30)$$

and

$$\phi(p) = \frac{\left\{\begin{array}{c} G(p)g(p)V_b \sin\delta_0(X_{dm}X_qV_q'V_b\sin\delta_0 \\ -X_{qm}X_dV_d'V_b\cos\delta_0) + V_b^2\sin^2\delta_0 X_q[X_d(p) - X_d]\end{array}\right\}}{X_dX_q[G(p)g(p)V_q'X_t - X_d(p)]} \quad . \quad . \quad . \quad (31)$$

Rewriting eqn. (28), we get

$$\frac{\Delta\delta}{\Delta T_i} = \frac{1/Mp^2}{1 + f(p)/Mp^2} \quad . \quad . \quad . \quad . \quad . \quad (32)$$

The simple block schematic which represents eqn. (32) is shown in Fig. 5. This is the basic closed-loop pattern for a synchronous generator for small oscillations. It is immediately obvious from the diagram that the function $f(p)/Mp^2$, in the denominator of eqn. (32), is the open-loop transfer function in operational form for a synchronous machine. The Nyquist criterion may be applied directly to it for $p = -j\infty$ to $+j\infty$ in order to determine whether the machine is stable and the degree of stability. In the following examples of the application of the frequency-response method, the principal parameters for the generator and transmission line are given in Table 1. Several parameters, such as generator and exciter time-constants and the regulator gain, are considered as variables except where otherwise stated.

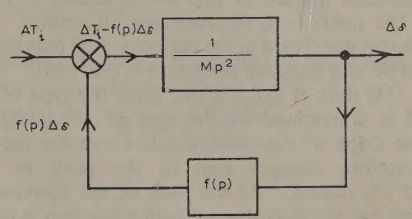


Fig. 5.—Basic closed-loop pattern for synchronous machine derived from small-displacement theory.

Table 1

$X_{dm} = 1{\cdot}0$ per unit.	$X'_{dm} = 0{\cdot}2$ per unit.
$X_{qm} = 0{\cdot}7$ per unit.	$H = 5$ kW s/kVA.
$X_t = 0{\cdot}5$ per unit.	$V_b = 1$ per unit.

(3.2.1) Synchronous Machine without Voltage Regulator.

In this and the following two Sections the initial condition of excitation corresponds to unit terminal voltage at unit power output.

The first example is a simple plot of the open-loop transfer function for an unregulated generator for several operating angles. The open-loop transfer function is

$$\frac{f(p)}{Mp^2} = \frac{1}{Mp^2}\left[\phi(\delta_0) + \frac{V_b^2\sin^2\delta_0(X_d - X_d')\tau_{d0}p}{X_d(X_d + X_d'\tau_{d0}p)}\right] \quad . \quad (33)$$

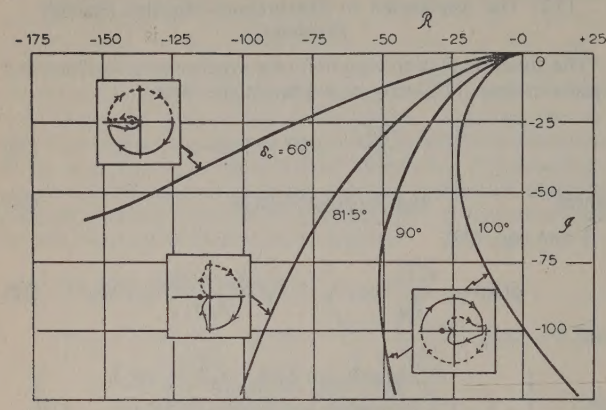


Fig. 6.—Frequency response of unregulated machine for various operating conditions.

This function is plotted in Fig. 6 and the nature of the closure of the curves in the region $p = j0$ is given with each curve. From eqn. (33) it may may be deduced that the open-loop transfer function has no poles with positive real parts, and therefore for stable operation there must be no net encirclement of the point $(-1, 0)$.

The significant feature of this set of curves is that, in the region of the point $(-1, 0)$, they are all similar. It is only when the curves are closed in the region $p = j0$ that the difference between stable and unstable operation is observed. It is evident from eqn. (33) that, if $\phi(\delta_0)$ is not zero, the type of closure to be applied is determined by the sign of $\phi(\delta_0)/M(j\omega)^2$. Now $\phi(\delta_0)$ is the slope of the power-angle curve for the generator, and it therefore changes sign as the peak in this curve ($\delta_0 = 81{\cdot}5°$) is passed. At the peak of the power-angle curve $\phi(\delta_0)$ is zero and the open-loop transfer function is proportional to $1/p$ in the region $p = j0$. Thus a third type of closure must be applied.

(3.2.2) Synchronous Machine with Voltage Regulator—Effect of Gain and Exciter Time-Constant.

The results in this Section appertain to the basic system of Fig. 1. Two important parameters of the regulator are the time-constant of the exciter and the regulator gain. In Fig. 7

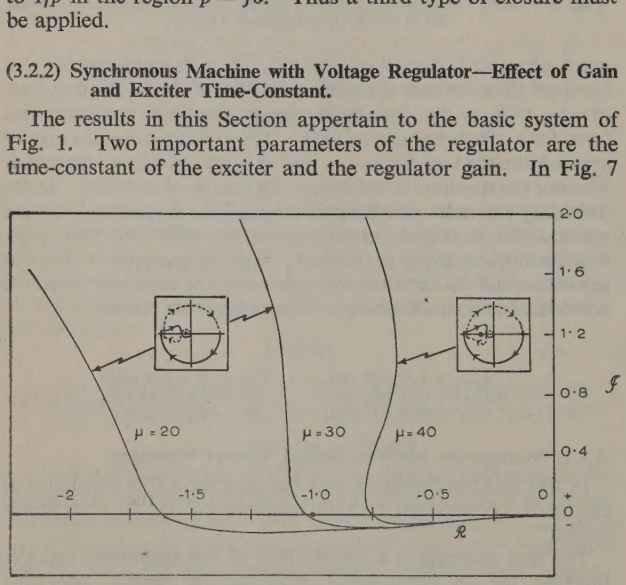


Fig. 7.—Frequency response of regulated generator showing effect of regulator gain.

$\delta_0 = 90°$ $K_d = 0$
$\tau_{d0} = 5$ sec $\tau_s = \mu_s = 0$
$\tau_e = 1$ sec $K = 0$

the function $f(j\omega)/M(j\omega)^2$ is plotted for three different values o μ. As μ is increased the function approaches the point $(-1, 0)$ For $\mu = 20$ and 30 the number of net encirclements is zero but for $\mu = 40$ there is one net encirclement. Since there ar no poles of the function $f(p)/Mp^2$ with positive real parts, th curve for $\mu = 40$ indicates that self-excited oscillations will occu

The influence of the exciter time-constant is illustrated in Fig. 8

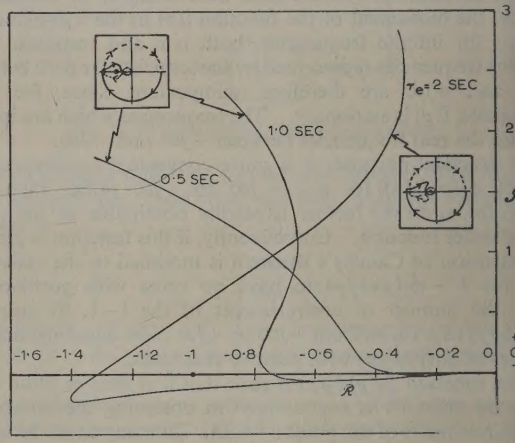


Fig. 8.—Frequency response of regulated generator showing effect of exciter-field time-constant.

$\delta_0 = 90°$ $\mu = 40$
$\tau_{d0} = 5$ sec $\tau_s = \mu_s = 0$
$K_d = 0$ $K = 0$

which shows that, for $\tau_e = 1{\cdot}0$ and $0{\cdot}5$ sec, the system unstable, whereas, for $\tau_e = 2$ sec, stability is achieved. Th may, at first, appear to be an unusual result, but it is verifie to some extent by the damping-coefficient characteristics Fig. 13. An increase of the exciter field time-constant resultin in greater stability has also been observed by Messerle an Bruck.[10]

(3.2.3) Influence of Positive Feedback Proportional to Rate of Chan of Field Current.

In References 11 and 6 a technique is described for improvin steady-state and transient-stability boundaries by including the regulator a subsidiary positive feedback proportional to th rate of change of field current. For the positive feedback have the desired effect it is necessary to nullify the exciter fie time-constant by a phase-advance circuit. In Section 4 it

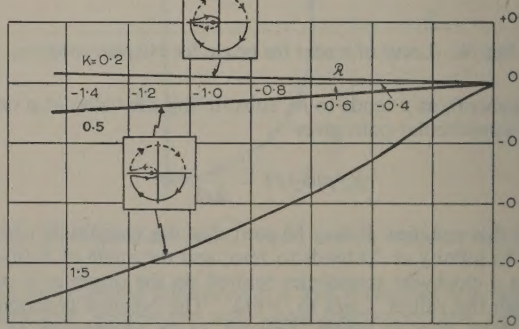


Fig. 9.—Frequency response of regulated machine showing effect o positive feedback.

$\tau_{d0} = 5$ $\delta_0 = 60°$
$\mu = 40$ $\mu_s = \tau_s = 0$

own that the positive feedback has marked effect on the mping in the system, depending on the magnitude of the feed-ck. It is interesting to observe the corresponding effect on Nyquist diagram, as shown in Fig. 9. For a small amount positive feedback ($K = 0{\cdot}2$) the system is unstable, the fre-ency characteristic showing one net encirclement of $(-1, 0)$. K is increased the frequency characteristics move away from point $(-1, 0)$ with zero net encirclements. Fig. 10 shows

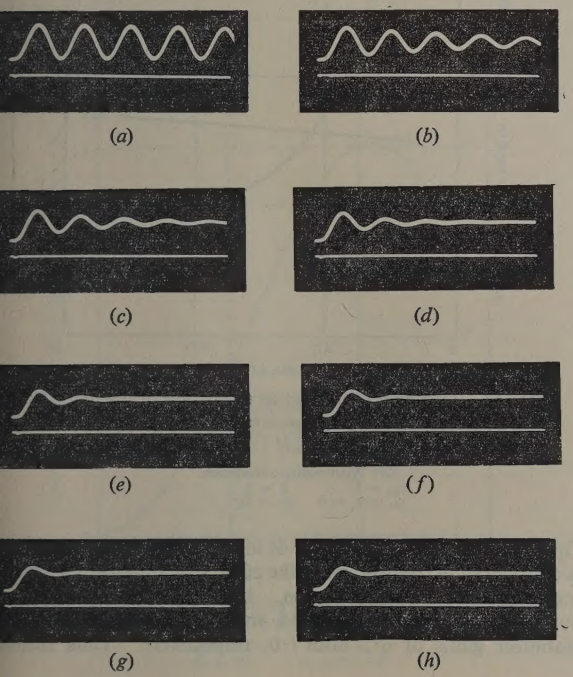


Fig. 10.—Effect of positive feedback on rotor-angle variation.

(*a*) $K = 0{\cdot}75$ (*e*) $K = 1{\cdot}75$
(*b*) $K = 1{\cdot}0$ (*f*) $K = 2{\cdot}0$
(*c*) $K = 1{\cdot}25$ (*g*) $K = 2{\cdot}25$
(*d*) $K = 1{\cdot}5$ (*h*) $K = 2{\cdot}5$

corresponding effect on transient stability as the positive dback is increased.

SOME ASPECTS OF DAMPING IN A SYNCHRONOUS GENERATOR

t was indicated in Section 2 that the damping inherent in system could be cancelled in the equations for small oscilla-ns by including an equal and opposite damping torque $D\Delta\delta$ in the mechanical equations of motion. The value of required to make the system oscillate with constant amplitude es a measure of the damping in the system for the given ial conditions and system parameters. A study of the mping is useful in that it throws some light on the mechanism tability and is also useful for purposes of comparison. The mping torque coefficients may be calculated from eqns. (24), and (26), or, if the equations of the system are set up on an logue computer, the values of K_d required for small oscilla-ns to exist may be obtained experimentally. The results sented in this Section were obtained by the latter procedure checked in many instances by mathematical computation. The generator and transmission-line parameters employed in se experiments are given in Table 1.

(4.1) Unregulated Generator

The initial operating point of the machine is chosen to be the rated terminal voltage at rated power output. In Fig. 11(*a*)

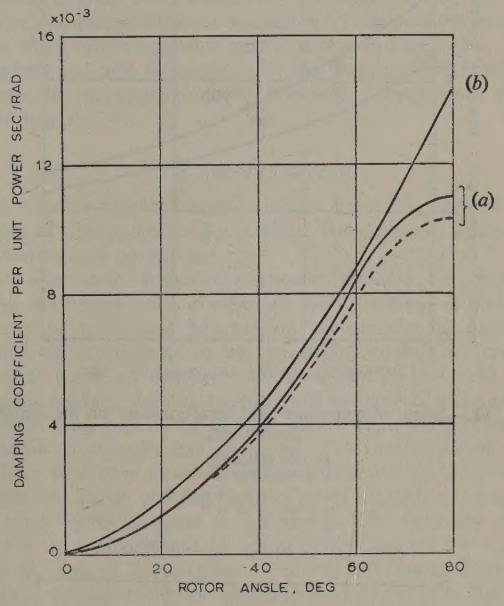


Fig. 11.—Influence of operating condition on the damping coefficient.

(*a*) No regulator.
——— Experimental.
– – – – Theoretical.
(*b*) With regulator.
——— Experimental.

$\tau_{d0} = 5$ $\mu = 40$ $K = 0$
$\tau_e = 2$ $\mu_s = \tau_s = 0$

the damping torque is plotted as a function of rotor angle. At $\delta_0 = 0$ the damping is zero and the machine hunts continuously. As the rotor angle increases so does the damping, and this is always positive.

The effect of the generator time-constant is shown in Fig. 12(*a*). From eqn. (24) the damping torque is seen to be zero both for infinite and zero field time-constants. Fig. 12(*a*) shows that, over the range of τ_{d0} considered, the damping torque decreases as τ_{d0} is increased. A maximum value for K_d exists between the values $\tau_{d0} = 0$ and 3.

(4.2) Generator with Voltage Regulator

The initial operating point at which the regulator is set up is again chosen to be the rated power output at rated terminal voltage. The first test in this study is to measure the damping coefficient as a function of rotor angle, and the result is shown in Fig. 11(*b*). The damping torque is observed to be slightly greater than that for the unregulated machine. The effect of the generator field time-constant is shown in Fig. 12(*b*). The damping coefficient is again slightly greater than that for the unregulated machine. Eqn. (25) indicates that the damping is zero both for infinite and zero generator-field time-constants. The curves of Fig. 13, showing the effect of exciter-field time-constant on the damping coefficient, demonstrate that the damping in the system may become negative for small exciter time-constants. This is an interesting result since it suggests that, for a finite τ_{d0}, there is a lower limit to τ_e for optimum damping.

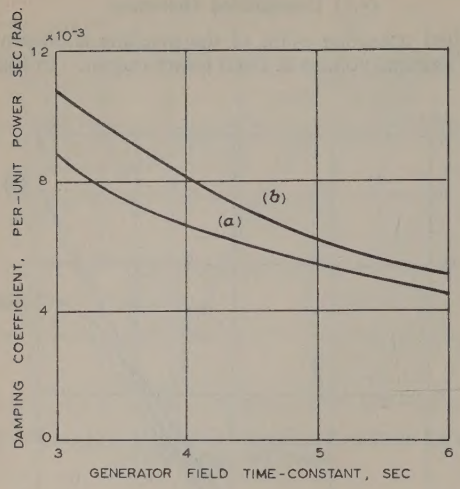


Fig. 12.—Effect of generator-field time-constant on the damping coefficient.

(*a*) No regulator.
(*b*) Regulator present.
$\tau_e = 2$ $\mu_g = \tau_g = 0$
$\mu = 40$ $K = 0$
$\delta_0 = 50°$

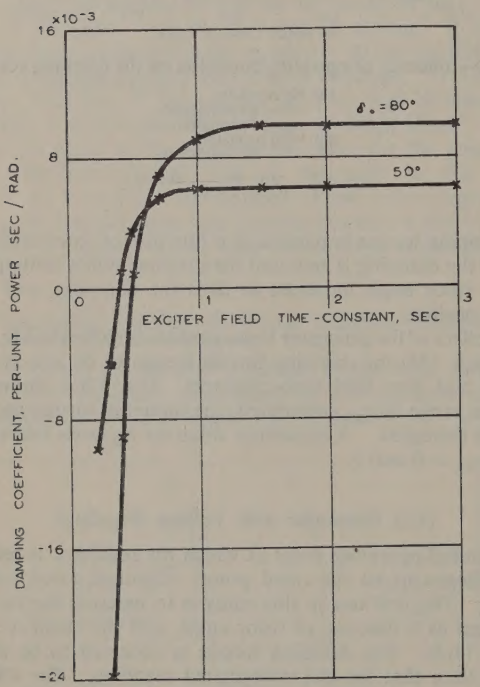


Fig. 13.—Damping coefficient as a function of exciter-field time-constant.

$\tau_{d0} = 5{\cdot}0$ $\mu_g = \tau_s = 0$
$\mu = 40$ $K = 0$

The curve of Fig. 14(*a*) shows how the regulator gain affects the damping. Changes in regulator gain appear to have little effect on the damping coefficient.

The effects of the gain and time-constant of a stabilizing transformer on the damping coefficient are shown in Fig. 15.

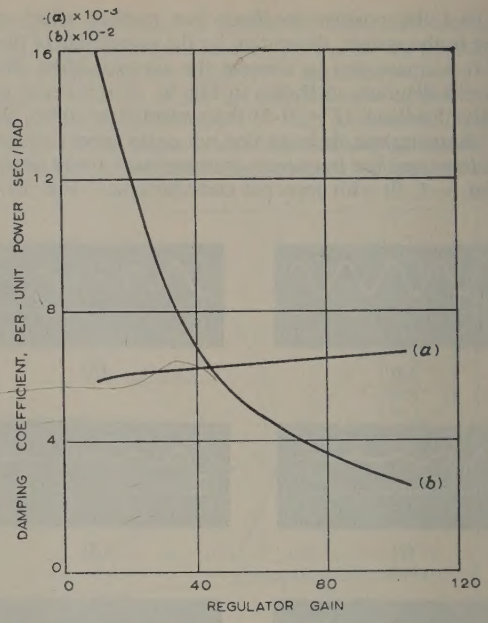


Fig. 14.—Damping coefficient as a function of regulator gain.

(*a*) Without positive feedback.
$\tau_{d0} = 5$ $\mu_g = \tau_g = 0$ $\delta_0 = 50°$
$\tau_e = 2$ $K = 0$
(*b*) With positive feedback.
$\tau_{d0} = 5$ $K = 2$
$\mu_g = \tau_g = 0$ $\delta_0 = 50°$

Operation at $\delta_0 = 90$ ($P_i = 1{\cdot}6$) in the dynamic stability regio is considered in this test, since the effect of a stabilizer is generall most noticeable in this region. It may be observed that fo curves (*b*) and (*c*) of Fig. 15 the damping changes sign a stabilizer gains of 6·5 and 7·0, respectively. Thus stabilize

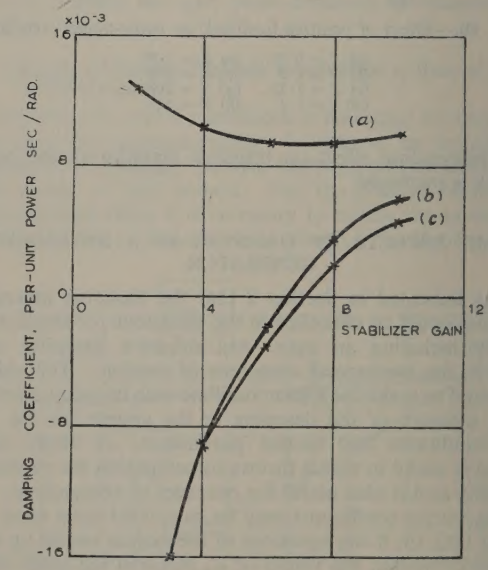


Fig. 15.—Influence of stabilizer gain and time-constant on the damping coefficient.

$\mu = 40$ $\delta_0 = 90°$
$K = 0$ $\tau_{d0} = 5$
(*a*) $\tau_e = 2$ $\tau_g = 2$
(*b*) $\tau_e = \frac{1}{4}$ $\tau_g = \frac{1}{4}$
(*c*) $\tau_e = \frac{1}{2}$ $\tau_g = 2$

n is a critical parameter for the correct functioning of a bilizing transformer.

3) Generator with Voltage Regulator incorporating a Positive Feedback Proportional to Rate of Change of Field Current

The positive feedback is identical with that described in ferences 11 and 6. The effect of a positive feedback on the nping coefficient for various rotor angles is shown in Fig. 16.

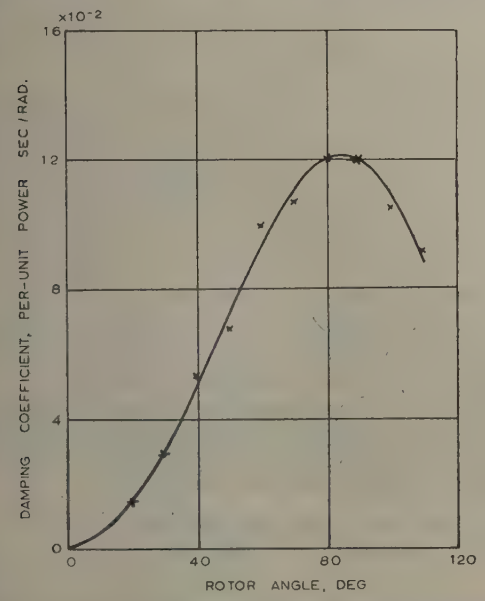


Fig. 16.—Damping coefficient as a function of rotor angle, with positive feedback.

$\tau_{d0} = 5{\cdot}0$ $K = 2{\cdot}0$
$\mu_r = 40$ $\mu_s = \tau_s = 0$

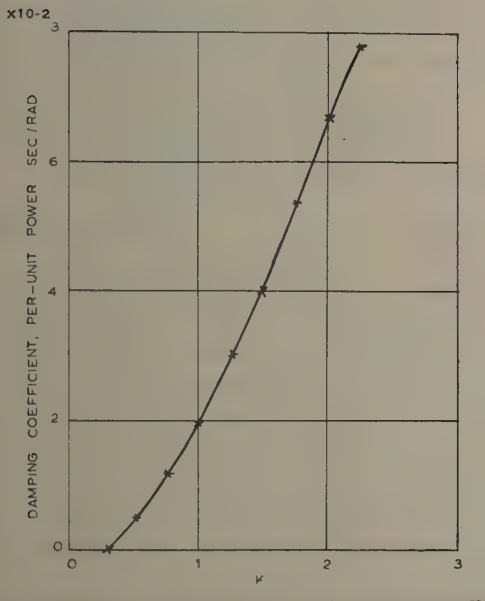


Fig. 17.—Effect of positive feedback on the damping coefficient.

$\tau_{d0} = 5{\cdot}0$ $\delta_0 = 50°$
$\mu_r = 40$ $\mu_s = \tau_s = 0$

By comparing this curve with that of Fig. 11, the positive-feedback loop is seen to bring about a very marked increase in the damping coefficient. Fig. 16 also shows that the damping coefficient is reduced as the load angle advances beyond its normal (i.e. with fixed excitation) maximum. Fig. 17 shows how the damping coefficient is affected by the positive feedback.

The effect of the regulator gain is shown in Fig. 14(*b*). From this graph and also from eqn. (26) the damping coefficient may be seen to be approximately inversely proportional to the regulator gain.

(5) CONCLUSIONS

The small-oscillation study reveals how the various components of the regulator loop affect the damping inherent in the synchronous generator.

The open-loop frequency-response method for predetermining mathematically whether or not a system is stable, in addition to providing information concerning its degree of stability, is considered to be an improvement on previous analytical work on synchronous machines in this field. The technique enables stability problems to be tackled with no more apparatus than pencil and paper. However, a word of caution is necessary to qualify this statement. The authors have stated elsewhere that they have experienced difficulty in inserting the appropriate initial conditions into small-oscillation equations. This observation is again emphasized. The difficulty can be obviated by measurement of initial conditions on the machine or by the use of a computer. Failing this the authors have found the voltage-excitation characteristics of Fig. 6 in Reference 6 to be most useful for determining the necessary initial conditions, and the work described has been framed around these characteristics.

The authors believe that the frequency-response method should prove of value in the design of damper windings, since the effect of variations in damper-winding resistances and reactances on stability may be conveniently observed on the Nyquist diagrams.

Finally, it should be noted that the small-oscillation method specifically restricts the changes in the system to be vanishingly small. Thus it is not strictly correct to interpret small-oscillation results as being directly applicable to conditions of large oscillations, although the authors believe that there is a correspondence between the two.

(6) ACKNOWLEDGMENTS

The authors are grateful to Prof. J. M. Meek for his interest in the work and for facilities provided in the laboratories of the Electrical Engineering Department of Liverpool University. They are also indebted to the Central Electricity Generating Board for financial support.

(7) REFERENCES

(1) CONCORDIA, C.: 'Steady-State Stability of Synchronous Machines as affected by Voltage-Regulator Characteristics', *Transactions of the American I.E.E.*, 1944, **63**, p. 215.

(2) ROUTH, E. J.: 'Advanced Rigid Dynamics' (Macmillan and Co.), p. 168.

(3) PARK, R. H.: 'Two-Reaction Theory of Synchronous Machines, Generalized Method of Analysis—Part II', *Transactions of the American I.E.E.*, 1933, **52**, p. 352.

(4) NYQUIST, H.: 'Regeneration Theory', *Bell System Technical Journal*, 1932, **2**, p. 126.

(5) ALDRED, A. S., and DOYLE, P. A.: 'Electronic-Analogue-Computer Study of Synchronous-Machine Transient Stability', *Proceedings I.E.E.*, Paper No. 2015 S, June, 1956 (**104** A, p. 152).
(6) ALDRED, A. S., and SHACKSHAFT, G.: 'The Effect of a Voltage Regulator on the Steady-State and Transient Stability of a Synchronous Generator', *ibid.*, Paper No. 2662 S, August, 1958 (**105** A, p. 420).
(7) CORLESS, K., and ALDRED, A. S.: 'An Experimental Electronic Power-System Simulator', *ibid.*, Paper No. 2673 S, October, 1958 (**105** A, p. 503).
(8) PARK, R. H.: 'Two-Reaction Theory of Synchronous Machines, Generalized Method of Analysis—Part I', *Transactions of the American I.E.E.*, 1929, **48**, p. 716.
(9) PARK, R. H.: 'Definition of an Ideal Synchronous Machine and Formula for the Armature Flux Linkages', *General Electric Review*, 1928, **31**, p. 332.
(10) MESSERLE, H. K., and BRUCK, R. W.: 'Steady-State Stability of Synchronous Generators as affected by Regulators and Governors', *Proceedings I.E.E.*, Monograph No. 134 S, June, 1955 (**103** C, p. 24).
(11) KRON, G.: 'A Super Regulator', *Matrix and Tensor Quarterly*, 1955, **5**, p. 71.

(8) APPENDIX

(8.1) Operational Equations of a Synchronous Machine and Voltage Regulator

The system to be analysed is shown in block schematic form in Fig. 1.

Park's[8] equations for an ideal synchronous machine,[9] with no amortisseur windings, connected to an infinite system via a transmission line are as follows:

Direct axis $V_{dm} = p\Phi_{dm} - \Phi_{qm}p\theta - R_m I_d$ . . (34)

Quadrature axis $V_{qm} = p\Phi_{qm} + \Phi_{dm}p\theta - R_m I_q$ . . (35)

Field $V_{fd} = p\Phi_{fd} + R_{fd}I_{fd}$ . . . . . (36)

in which $\Phi_{fd} = X_{fd}I_{fd} - X_{ad}I_d$ . . . . . (37)

$\Phi_{dm} = X_{ad}I_{fd} - X_{dm}I_d$ . . . . . (38)

$\Phi_{qm} = -X_{qm}I_q$ . . . . . (39)

The generated torque is

$$T_u = \Phi_{dm}I_q - \Phi_{qm}I_d \quad . . . . . (40)$$

The equations for a transmission line in Park's reference frame are as follows:

$$V_d = p\Phi_{dt} - \Phi_{qt}p\theta - R_t I_d + V_{dm} \quad . . . (41)$$

$$V_q = p\Phi_{qt} + \Phi_{dt}p\theta - R_t I_q + V_{qm} \quad . . . (42)$$

in which $\Phi_{dt} = -X_t I_d$ . . . . . (43)

$\Phi_{qt} = -X_t I_q$ . . . . . (44)

The suffixes m and t indicate machine and transmission line quantities respectively. The d and q equations for the machine may be added to the d and q equations for the transmission lin to give

$$V_d = p\Phi_d - \Phi_q p\theta - RI_d \quad . . . . (4$$

$$V_q = p\Phi_q + \Phi_d p\theta - RI_q \quad . . . . (4$$

in which $\Phi_d = \Phi_{dm} + \Phi_{dt}$ . . . . . (4

$\Phi_q = \Phi_{qm} + \Phi_{qt}$ . . . . . (4

and $R = R_m + R_t$ . . . . . (4

The direct- and quadrature-axis components of the busb voltage are as follows:

$$V_d = V_b \sin\delta \quad . . . . . (5$$

$$V_q = V_b \cos\delta \quad . . . . . (5$$

The equation of motion of the system is

$$T_i = M\frac{d^2\delta}{dt^2} + K_d\frac{d\delta}{dt} + \Phi_d I_q - \Phi_q I_d \quad . . (5$$

The following assumptions are also made:

(*a*) The speed change during a disturbance, which results in th machine remaining in synchronism, is a negligible fraction of funda mental speed, i.e. $p\theta$ is assumed constant and in the per-unit syste it is equal to unity.
(*b*) Armature and transmission-line resistance is negligible.
(*c*) The voltages induced in the armature by the rate of chang of armature flux linkages, namely $p\Phi_d$ and $p\Phi_q$, are negligible com pared with the voltages, $\Phi_q p\theta$ and $\Phi_d p\theta$, generated by the flux Φ_d and Φ_q rotating at fundamental speed.

With these assumptions the equations of a synchronou machine and transmission line reduce to

$$V_d = X_q I_q \quad . . . . . (53$$

$$V_q = G(p)V_f - X_d(p)I_d \quad . . . . . (54$$

$$V_d = V_b \sin\delta \quad . . . . . (55$$

$$V_q = V_b \cos\delta \quad . . . . . (56$$

$$T_i = Mp^2\delta + K_d p\delta + \Phi_d I_q - \Phi_q I_d \quad . . (57$$

The operational expressions for $G(p)$ and $X_d(p)$ are as follows

$$G(p) = \frac{1}{1 + \tau_{d0}p} \quad . . . . . (58$$

$$X_d(p) = \frac{X_d + X'_d\tau_{d0}p}{1 + \tau_{d0}p} \quad . . . . (59$$

For the voltage regulator in Fig. 1 the equations are as follows

$$V_f = g(p)[V_m - V_r] \quad . . . . . (60$$

$$V_m^2 = V_{dm}^2 + V_{qm}^2 \quad . . . . . (61$$

$$V_{dm} = V_d - X_t I_q \quad . . . . . (62$$

$$V_{qm} = V_q + X_t I_d \quad . . . . . (63$$

where $g(p) = \frac{-\mu}{1 + \tau_e p}$ . . . . . (64

621.313.013.5:621.317.79

The Institution of Electrical Engineers
Monograph No. 341 M
July 1959
©

ANALOGUE TREATMENT OF EDDY-CURRENT PROBLEMS INVOLVING TWO-DIMENSIONAL FIELDS

By J. ROBERTS, B.Sc.(Eng.), Associate Member.

(*The paper was first received* 18*th December*, 1958, *and in revised form* 22*nd April*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in* July, 1959.)

SUMMARY

The paper deals with the *RC*-analogue representation of electromagnetic fields involving unidirectional current flow, such as those associated with current-carrying conductors lying in armature slots. The resulting magnetic field is in general 2-dimensional and is fully represented in the analogue. Capacitor current in the latter is directly equivalent to current flow in the corresponding section of the conductor, and the eddy-current loss in a conductor may be calculated from the distribution of current between the several capacitors. Some simple examples illustrate the translation of field boundary conditions into analogue terms.

LIST OF SYMBOLS

E, E_z = Electric field strength, V/m.
H, H_x, H_y = Magnetic field strength, A/m.
J, J_z = Current density, A/m^2.
i = Total current in a single conductor, amp.
Φ = Magnetic flux per unit length in z-direction, Wb/m.
σ = Conductivity, mhos/m.
μ = Permeability, H/m.
$\omega/2\pi$ = Frequency, c/s.
$\delta = \sqrt{(2/\omega\mu\sigma)}$ = Skin depth in conductor, m.
R_{DC} = Resistance of conductor by d.c. measurement, ohms.
R_{AC} = Effective a.c. resistance due to eddy currents, ohms.
h = Interval in finite difference approximation, m.
K = Scaling factor, m/s.
R = Analogue resistance, ohms.
C = Analogue capacitance, F.
I_x, I_y = Analogue resistor currents, amp.
I_0, I_1, . . . I_r = Analogue capacitor currents, amp.
V = Potential in resistor plane, volts.
V' = Potential of common capacitor junction in conductor representation, volts.

(1) INTRODUCTION

The non-uniform distribution of current in conductors embedded in an armature slot has been determined analytically for a simple geometrical arrangement by several authors.[1–3] Provided that the slot has a rectangular cross-section of small width/depth ratio and that the conductors are of rectangular cross-section and substantially fill the width of the slot but do not approach closely its open end, it is a fair assumption that the magnetic flux paths are straight lines normal to the sides of the slot. Under these conditions the current density is a function of only one space co-ordinate, namely depth in the slot, and its evaluation is a straightforward solution of the 1-dimensional diffusion equation, with appropriate boundary conditions.

Recent trends in machine design have produced several examples of conductor and slot geometry which do not fulfil these conditions. For instance, an outstanding advance has been made in respect of increased power rating of machines of given overall dimensions by the introduction of direct cooling of hollow armature conductors. The magnetic flux pattern in such a case is not likely to be as simple as that assumed in the 1-dimensional case, and furthermore the current density will be influenced by the variation of effective conductor width with depth in the slot. Even so, the flow of current is along the length of the conductors and the magnetic field lines are confined to planes normal to the conductor length. The field equations in this case are particularly suited to a simple analogue treatment,[4,5] and the current density in the conductors has a counterpart in the analogue which may conveniently be measured.

(2) FIELD EQUATIONS FOR UNIDIRECTIONAL CURRENT FLOW

The field equations will be set up here for a continuous medium of conductivity σ and permeability μ, taking account of the simplifications appropriate to this particular problem. Assuming sinusoidal time variation of all field quantities at a frequency $\omega/2\pi$, and ignoring displacement current, which is not significant in this case, Maxwell's equations take the form:

$$\text{curl } H = J \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$\text{curl } E = -j\omega\mu H \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

J and E being related by

$$J = \sigma E$$

The time factor, $\exp j\omega t$, is omitted for ease of writing.

Choosing the z-axis as the direction of current flow in the medium, J and E will only have components J_z and E_z. Hence eqns. (1) and (2), in component form, reduce to:

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = J_z \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

$$\frac{\partial J_z}{\partial y} = -j\omega\mu\sigma H_x \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$\frac{\partial J_z}{\partial x} = j\omega\mu\sigma H_y \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

It is clear that, with the assumption of unidirectional current density, J_z, there is no z-component of magnetic field strength and no z-dependence of any field quantity, i.e. the field is essentially 2-dimensional. Elimination of H_x and H_y from eqns. (3), (4) and (5) shows that J_z must satisfy

$$\frac{\partial^2 J_z}{\partial x^2} + \frac{\partial^2 J_z}{\partial y^2} = j\omega\mu\sigma J_z \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Correspondence on Monographs is invited for consideration with a view to publication.
Mr. Roberts is at the Imperial College of Science and Technology, University of London.

(2.1) In Finite-Difference Form

Consider the xy-plane in the continuous medium divided into square elements, the length of side h being so small that h^2 multiplied by the current density at the centre of an element is a good approximation to the total current flowing through this element. Fig. 1 shows one such element whose centre-point,

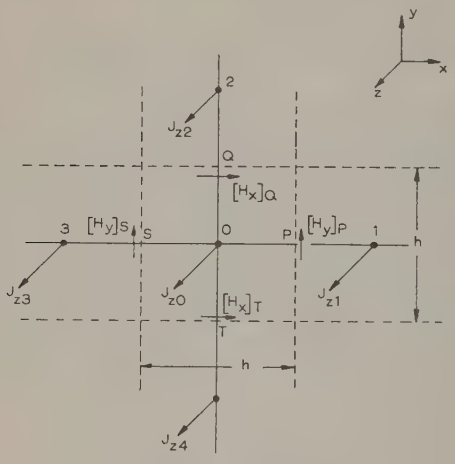


Fig. 1.—Cross-section of continuous medium perpendicular to the direction of current flow.

Dotted lines mark the boundary of a square element of side h used in the finite-difference approximation.

0, is the origin of the x- and y-axes. Points 1, 2, 3 and 4 are the centre-points of neighbouring elements. Using Maclaurin's theorem, the current densities at points 1 and 3 may be expressed in terms of J_z and its derivatives at 0 as follows:

$$J_{z1} = J_{z0} + h\left[\frac{\partial J_z}{\partial x}\right]_0 + \frac{h^2}{2}\left[\frac{\partial^2 J_z}{\partial x^2}\right]_0 + \frac{h^3}{6}\left[\frac{\partial^3 J_z}{\partial x^3}\right]_0 + \frac{h^4}{24}\left[\frac{\partial^4 J_z}{\partial x^4}\right]_0 + \ldots \quad . \quad . \quad . \quad . \quad (7a)$$

and $$J_{z3} = J_{z0} - h\left[\frac{\partial J_z}{\partial x}\right]_0 + \frac{h^2}{2}\left[\frac{\partial^2 J_z}{\partial x^2}\right]_0 - \frac{h^3}{6}\left[\frac{\partial^3 J_z}{\partial x^3}\right]_0 + \frac{h^4}{24}\left[\frac{\partial^4 J_z}{\partial x^4}\right]_0 - \ldots \quad . \quad . \quad . \quad . \quad (7b)$$

where the numerical subscripts identify the location at which the field quantities are applicable.

Adding eqns. (7a) and (7b),

$$J_{z1} + J_{z3} = 2J_{z0} + h^2\left[\frac{\partial^2 J_z}{\partial x^2}\right]_0$$

provided that h is so small that $\frac{h^4}{12}\left[\frac{\partial^4 J_z}{\partial x^4}\right]_0$ is negligible.

Hence $$\left[\frac{\partial^2 J_z}{\partial x^2}\right]_0 = (J_{z1} + J_{z3} - 2J_{z0})/h^2 \quad . \quad . \quad . \quad (8a)$$

Similarly $$\left[\frac{\partial^2 J_z}{\partial y^2}\right]_0 = (J_{z2} + J_{z4} - 2J_{z0})/h^2 \quad . \quad . \quad . \quad (8b)$$

Replacing the second derivatives in eqn. (6) by the finite-difference approximations of eqns. (8a) and (8b) gives

$$J_{z1} + J_{z2} + J_{z3} + J_{z4} - 4J_{z0} = j\omega\mu\sigma h^2 J_{z0} \quad . \quad . \quad (9)$$

The first space derivatives of the field quantities may be approximated in the same way.

Thus $$\left[\frac{\partial H_y}{\partial x}\right]_0 = \{[H_y]_P - [H_y]_S\}/h, \text{ etc.}$$

and $$\left[\frac{\partial J_z}{\partial x}\right]_P = (J_{z1} - J_{z0})/h, \text{ etc.}$$

where the subscripts P, Q, S and T refer to the values of the field quantities at the centre-points of the sides of the element as shown in Fig. 1. Eqns. (3), (4) and (5) may therefore be expressed in the following finite-difference forms:

$$h^2 J_{z0} = h\{[H_y]_P - [H_y]_S + [H_x]_T - [H_x]_Q\} \quad . \quad (10)$$

$$\mu h[H_x]_Q = (h^2 J_{z0} - h^2 J_{z2})/j\omega\sigma h^2 \quad . \quad . \quad (11a)$$

$$\mu h[H_x]_T = (h^2 J_{z4} - h^2 J_{z0})/j\omega\sigma h^2 \quad . \quad . \quad (11b)$$

$$\mu h[H_y]_P = (h^2 J_{z1} - h^2 J_{z0})/j\omega\sigma h^2 \quad . \quad . \quad (12a)$$

$$\mu h[H_y]_S = (h^2 J_{z0} - h^2 J_{z3})/j\omega\sigma h^2 \quad . \quad . \quad (12b)$$

(3) EQUIVALENT NETWORK

Fig. 2 shows a portion of a network consisting of equal resistors, R, connected in a square mesh pattern and lying in a

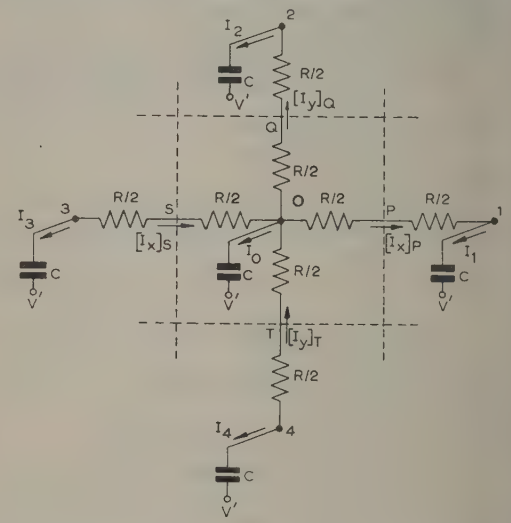


Fig. 2.—Portion of a network analogue for a continuous conducting medium.

single plane. From each resistor junction a capacitor, C, is returned to a common point of alternating potential V'. In the diagram the resistors are divided into units of $R/2$, since it is intended to show the equivalence of the element of the network delineated by the dotted line to the square element of side h in Fig. 1. Alternating voltages are applied to the network at some remote points, as yet unspecified, giving rise to the distribution of voltages and currents indicated in the diagram.

The following relationships between the voltages and currents of Fig. 2 are apparent:

$$\left.\begin{aligned} V_0 - V_1 &= [I_x]_P R \\ V_0 - V_2 &= [I_y]_Q R \\ V_3 - V_0 &= [I_x]_S R \\ V_4 - V_0 &= [I_y]_T R \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (13)$$

Also
$$\left.\begin{aligned} I_0 &= j\omega C(V_0 - V') \\ I_1 &= j\omega C(V_1 - V') \\ I_2 &= j\omega C(V_2 - V') \\ I_3 &= j\omega C(V_3 - V') \\ I_4 &= j\omega C(V_4 - V') \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (14)$$

Equating the currents at point 0,

$$\begin{aligned} I_0 &= -[I_x]_P - [I_y]_Q + [I_x]_S + [I_y]_T \\ &= (V_1 + V_2 + V_3 + V_4 - 4V_0)/R \text{ [from eqn. (13)]} \\ &= (I_1 + I_2 + I_3 + I_4 - 4I_0)/j\omega CR \text{ [from eqn. (14)]} \end{aligned}$$

i.e.
$$I_1 + I_2 + I_3 + I_4 - 4I_0 = j\omega CRI_0 \quad . \quad . \quad . \quad (15)$$

Comparing eqns. (9) and (15) it is seen that the capacitor current in the network satisfies the same difference equation as the current density in the original field, and the network is a true analogue, provided that

$$\omega CR = \omega\mu\sigma h^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

Multiplication of eqn. (9) throughout by h^2, the area of each element of the field, shows that the total current per element also satisfies the difference equation. The capacitor current will therefore be taken to represent the total current crossing the corresponding elemental area by putting

$$I_0 = h^2 J_{z0} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

Expressing I_0 in terms of analogue potentials, from eqn. (14),

$$j\omega C(V_0 - V') = h^2\sigma E_{z0}$$

By setting up
$$C = h^2\sigma/K \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

where K is some convenient scaling factor carrying the dimensions metres per second,

$$V_0 - V' = KE_{z0}/j\omega \quad . \quad . \quad . \quad . \quad . \quad (19)$$

The choice of R is now settled by eqns. (16) and (18),

i.e.
$$R = K\mu \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

From eqns. (13) and (14),

$$R[I_y]_Q = V_0 - V_2 = (I_0 - I_2)/j\omega C$$

Comparing this with eqn. (11*a*), it is clear that the foregoing relationships produce the following equivalence:

$$I_y = hH_x \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (21a)$$

Similarly
$$I_x = -hH_y \quad . \quad . \quad . \quad . \quad . \quad (21b)$$

Several general features of the analogue may be noted from these equations:

(*a*) Resistance and capacitance in the analogue correspond to permeability and conductivity respectively in the medium.

(*b*) From eqns. (21*a*) and (21*b*) it is clear that a resistor current indicates the line integral of magnetic field strength, or m.m.f., in the xy-plane along the side of the element normal to the resistor length.

(*c*) Eqn. (19) shows that the potential difference across a capacitor is a measure of the electric field strength at the corresponding point in the medium, the phase quadrature arising as a result of the choice of analogue components.

(*d*) From eqn. (13),

$$\begin{aligned} V_1 - V_0 &= -[I_x]_P R \\ &= K\mu[H_y]_P h \text{ [from eqn. (21}b\text{)]} \\ &= K\Phi_{10} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (22) \end{aligned}$$

where Φ_{10} is the magnetic flux swept out per metre displacement in the z-direction of the line joining points 0 and 1. Consequently points in the resistance plane at the same potential define an area in the medium having zero normal component of magnetic flux, i.e. equipotential lines in the analogue correspond to magnetic flux lines in the original field.

(4) BOUNDARY CONDITIONS

Since the foregoing analysis has dealt specifically with a continuous medium, the first point which requires attention is the means and justification for connecting the analogue representations of separate media at a common interface. This will be illustrated by reference to Fig. 3, which shows the analogue

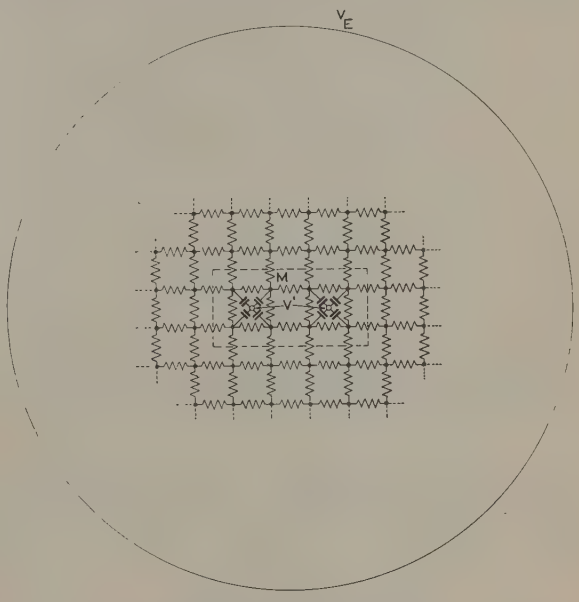


Fig. 3.—Network analogue for long straight copper conductor, of rectangular cross-section, in air.

Dotted line denotes the conductor boundary.

simulating a long straight copper conductor, of rectangular cross-section, in air. The copper is represented by the mesh enclosed by the dotted line, a suitable interval h being chosen bearing in mind the finite-difference approximation, and K being so chosen that reasonable values are obtained for R ($= K\mu$) and C ($= h^2\sigma/K$). Since air is non-conducting, no capacitors appear in the circuit outside the dotted line, and the value of the resistors is $K\mu_0$, where μ_0 is the free-space permeability constant. Note that the resistors here have the same value throughout the network since the relative permeability of copper is unity.

The direct connection of the two parts of the analogue at the dotted line imposes two conditions on the network. In the first place the potentials at the dotted line have been equated, which, in terms of the original field, implies equating the E_z values for the two media at the interface, since K has the same value throughout the analogue [compare eqn. (19)]. Secondly, the continuity of current flow through the resistors crossing the dotted line effectively equates the tangential components of magnetic field strength at the interface, as seen in eqn. (21), for the same interval h in the two media. The direct connection is thus justified, since the continuity of the tangential components of both the electric and magnetic field strengths at an interface is a well-known property of the electromagnetic field.

To show the way in which a.c. excitation is applied to the analogue the specific case will be considered of the conductor in Fig. 3 carrying a total current i. From the initial premise of eqn. (17), i is the total current leaving the network at the common capacitor junction; this junction is consequently one of the supply points of the analogue and has an indicated potential V'. To get a complete representation of the field due to this current-

carrying conductor, the current will have to be supplied to the analogue at points infinitely remote from the conductor. It is clear, however, that this current can equally well be supplied to the analogue by joining together points lying on any equipotential and raising this common lead to a potential V_E sufficient to produce the current i, provided that this equipotential encloses completely all parts of the field of interest. In particular, if only the current distribution in the conductor is required, any equipotential which lies wholly outside the dotted line may be chosen: in field terms, this means that the magnetic flux external to a conductor plays no part in determining the current distribution within it. Since equipotentials in the analogue correspond to lines of magnetic flux in the original field, the choice of an equipotential requires prior knowledge of the flux pattern in that region of the field. In the case of the isolated conductor, the lines of magnetic flux will approximate to circles as the distance from the conductor increases. It would therefore be reasonable here to choose circles of increasing radii until further increase produced no apparent change in the region of interest.

To examine further the significance of potential division within the network, suppose that the analogue is infinite in extent and that an alternating potential V_∞ is applied to infinitely remote points of the network. Under this excitation let the potential of some mesh point M within the dotted line be V_M. Integrating eqn. (22) from M to infinity,

$$V_\infty - V_M = K\Phi_M$$

where Φ_M is the total flux interlinking unit length of the z-directed line through the corresponding point M of the conductor cross-section. Hence

$$j\omega(V_\infty - V_M)/K = j\omega\Phi_M$$

i.e. the inductive voltage drop per unit length of the conductor at M. Also, if I_M is the current through the capacitor connected to mesh point M, and J_{zM} and E_{zM} are the current density and electric field strength at point M in the conductor,

$$\begin{aligned} j\omega(V_M - V')/K &= I_M/KC \\ &= h^2 J_{zM}\frac{1}{K}\frac{K}{h^2\sigma} \\ &= E_{zM} \end{aligned}$$

which is the resistive voltage drop per unit length of the conductor at M.

It is seen that $j\omega(V_\infty - V')/K$ is therefore the total applied voltage per unit length of the conductor. In the case considered here of a single conductor, no significance attaches to V', since the analogue currents depend solely on the total applied voltage $V_\infty - V'$. However, if several conductors are present it will be necessary in general to raise the common capacitor junctions of the separate conductors to different potentials (i.e. different values of V'), since the applied voltage per unit length will not usually be the same for each conductor. If the applied voltage per unit length is the same for each conductor, V' will be the potential of all the capacitor junctions: this clearly corresponds to a parallel connection of the conductors, in which case the separate conductors can be regarded simply as parts of a single conductor.

(5) CONDUCTORS IN AN ARMATURE SLOT

Fig. 4 shows an arrangement of four conductors in an armature slot (*a*) and the corresponding analogue (*b*). Here there is the additional complication of an iron circuit. The dotted lines in Fig. 4(*a*) representing lines of magnetic flux are intended to convey only the flux curvature near the mouth of the slot, the symmetry of flux distribution about the vertical centre line of the slot

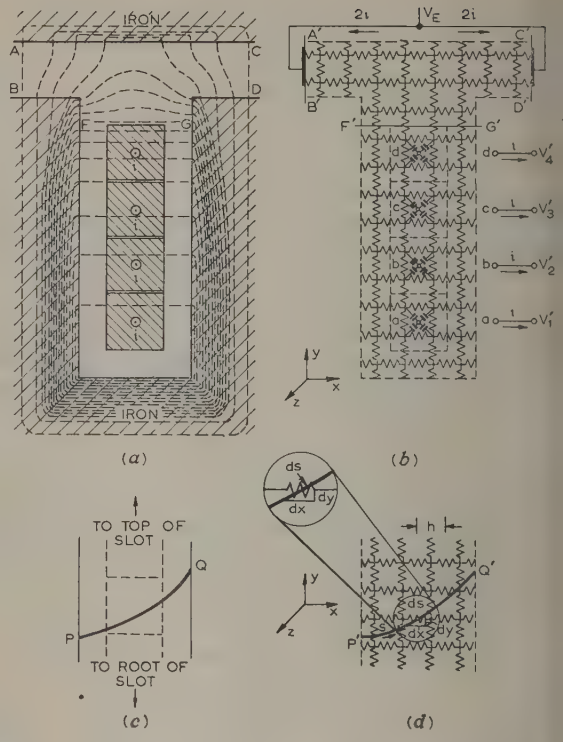


Fig. 4.—Analogue representation of current-carrying conductors an armature slot.

(*a*) Cross-sectional view showing arrangement of four conductors in a slot.
---- Lines of magnetic flux.
(*b*) Network analogue.
(*c*) Arbitrary line, PQ, across slot.
(*d*) Corresponding line, P′Q′, in analogue. Capacitors connected at heavily dot points are omitted for clarity.

and a pictorial indication of flux closure through the iron circu In the analogue representation the usual assumption of infin permeability of the iron has been made, leading to resistors infinite value in the analogue, i.e. the iron face is an open-circuit boundary. Expressed differently, there is no tangential magne field at the interface and consequently no current flow perpen cular to the boundary in the analogue.

If the magnetic flux paths in the iron are of interest, representation of this part of the field for a solid-iron armatu will follow the lines indicated in previous Sections. If iron is laminated, as is usual, with the laminations ly in planes z = constant, then current flow in the z-directi is suppressed and the field quantities are not strictly invaria with z. However, with effective lamination, the flux dens in the iron will not vary appreciably with z, and the sha fall-off of flux density in the intervening insulating layers maximum thickness 0·002 in) will not be apparent in the s at distances from the slot wall greater than the thickness of insulating layer approximately. In such a case an adequate rep sentation in the analogue is a resistor network without capaci loading, the value of the resistors being K times the avera weighted permeability of the laminations and intervening in lating layers.

The insulating layers between the conductors in the slot assumed to be of negligible thickness. Their sole function isolating the conductors is then catered for by isolating common capacitor junctions of the several conductors. T

continuity of tangential magnetic field strength through the layer is represented by the continuity of current flow through the resistors normal to the boundaries in the analogue.

As an example of setting up boundary conditions in this arrangement, the case will be considered in which the conductors are fully transposed, i.e. each conductor occupies each of the four possible positions in the slot over an equal length. This means that each conductor carries the same current i. Consequently, in the analogue the remote lines A′B′ and C′D′ are raised to a potential V_E, and V_1', V_2', V_3' and V_4' are so adjusted that equal currents leave the network through these connections. The lines A′B′ and C′D′ are chosen symmetrically disposed with respect to the slot, since their counterparts AB and CD in the original field must lie on the same closed line of magnetic flux; moreover, A′B′ and C′D′ are sufficiently remote from the slot that they can be judged with reasonable accuracy. If the top conductor is so far from the mouth of the slot that it does not enter the field fringing region, a single line of application of the voltage V_E such as F′G′ can be used.

As a check on the validity of the representation and, in particular, that the current distribution throughout the conductors is truly portrayed in the analogue, consider any arbitrary line PQ across the slot and its counterpart P′Q′ in the model [Fig. 4(*c*) and (*d*)]. Any elementary length ds of P′Q′ may or may not intercept resistors carrying current in the analogue. The probability of its intercepting an x-directed resistor is dy/h, and of its intercepting a y-directed resistor is dx/h. The mean current crossing such an element towards the root of the slot is therefore

$$I_x\frac{dy}{h} - I_y\frac{dx}{h}$$

Hence the total current crossing P′Q′, i.e. the total current leaving the analogue through capacitors connected on the root side of P′Q′, is

$$\int_{P'Q'}\left(\frac{I_x dy}{h} - \frac{I_y dx}{h}\right)$$

$$= -\int_{PQ}(H_y dy + H_x dx) \text{ [from eqn. (21)]}$$

$$= -\int_{PQ} \boldsymbol{H}\,.\,d\boldsymbol{s}$$

$$= \int_{QP} \boldsymbol{H}\,.\,d\boldsymbol{s}$$

the line integral of magnetic field strength from Q to P.

Since the magnetic circuit is completed through a zero-reluctance path around the slot root, this line integral is equal to the total z-directed current flowing through conductors, or parts of conductors, lying on the root side of PQ. The choice of PQ has been quite arbitrary; hence the equivalence of conductor currents and capacitor currents throughout the conductor region has been established.

(5.1) Calculation of Eddy-Current Losses and Effective Resistance

By measuring the current through individual capacitors, the current flow through the corresponding elements, of area h^2, of the conductors is determined. Fixing attention on one particular conductor, let its cross-section be represented by N square elements in the analogue. In the absence of eddy currents, the ohmic loss per metre in this conductor is given by

$$\tfrac{1}{2}\hat{\imath}^2 R_{DC} = \tfrac{1}{2}\hat{\imath}^2/N\sigma h^2$$

where R_{DC} is the value of resistance per metre obtained by d.c. measurement, and $\hat{\imath}$ is the peak value of the total conductor current, i, given by $i = \sum_{r=1}^{r=N} I_r$, where I_r is the current through the rth capacitor associated with this conductor.

The resistance per metre offered by the conductor to the current through each element is $1/\sigma h^2$.

$$\text{Hence, the actual power loss} = \sum_{r=1}^{r=N} \tfrac{1}{2}\hat{I}_r^2/\sigma h^2$$

$$= \frac{1}{2\sigma h^2}\sum_{r=1}^{r=N}\hat{I}_r^2$$

If this power loss be written in terms of an eff ective a.c. resistance R_{AC}, as $\frac{1}{2}\hat{\imath}^2 R_{AC}$,

$$\text{then} \qquad \frac{R_{AC}}{R_{DC}} = \frac{N\sum_{r=1}^{r=N}\hat{I}_r^2}{\left[\sum_{r=1}^{r=N} I_r\right]^2_{peak\ value}} \qquad . \quad . \quad . \quad . \quad (23)$$

(6) COMPARISON OF UNIDIMENSIONAL ANALOGUE RESULTS WITH ANALYTICAL FIELD SOLUTIONS

Although the value of this analogue technique rests mainly on its applicability to problems in which the time-varying magnetic field is 2-dimensional, a few 1-dimensional problems will now be considered, partly to show the use of different boundary conditions and how they are translated into analogue terms, and also by comparing the solutions with known analytical solutions to give an indication of reasonable values of the interval h.

(6.1) Semi-Infinite Solid

The eddy-current loss and flux penetration will be determined for a semi-infinite solid conductor, or infinitely thick plate, when it is subjected to an alternating magnetic field tangential to the surface. The surface is taken as the plane $x = 0$, and the y-axis as the direction of the magnetic field of strength H_{y0}.

A well-known result of the analytical solution of this problem is that the current density and magnetic field strength decrease with increasing depth in the solid, being reduced by a factor $1/\varepsilon$ in a distance $\delta = \sqrt{(2/\omega\mu\sigma)}$, known as the skin depth or depth of penetration. If the interval h chosen for the analogue is expressed in terms of δ, a solution is obtained which is applicable to any conducting solid. Let $h = \delta/2$: then, from eqn. (16),

$$\omega CR = \omega\mu\sigma h^2 = \tfrac{1}{2} \qquad . \quad . \quad . \quad . \quad . \quad (24)$$

It will be found unnecessary to specify C and R separately.

The analogue representation of this problem simplifies to a ladder network of infinite length, as shown in Fig. 5, since there

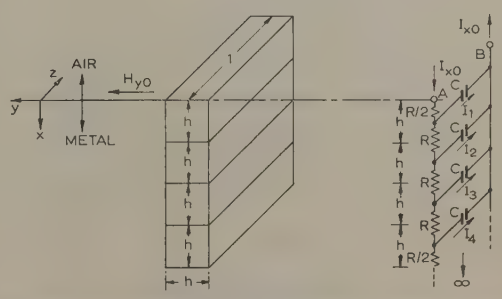


Fig. 5.—Network analogue for a semi-infinite conductor subjected to a uniform alternating magnetic field tangential to its surface.

is no y-variation, and each network section refers to an element of the conductor of cross-section $h \times h$ in the xy-plane. To represent the field in the conductor fully the voltage should be applied to the network between terminal A and an infinitely remote point along the resistor chain, but since no current will flow through capacitors at infinite values of x, the remote potential is that of terminal B and the voltage is consequently applied across AB. The given boundary condition in this problem is the value H_{y0} of the magnetic field strength at $x = 0$. Since, from eqn. (21b),

$$I_{x0} = -H_{y0}h = -H_{y0}\delta/2 \quad . \quad . \quad . \quad . \quad (25)$$

the applied voltage must be so adjusted that the input current, I_{x0}, to the line network is given by eqn. (25).

Analysis of this circuit shows that

$$\frac{I_{r+1}}{I_r} = \frac{1}{\left[j\omega CR - \frac{(\omega CR)^2}{4}\right]^{1/2} + \frac{j\omega CR}{2} + 1} = 0{\cdot}531 - j0{\cdot}282 \quad \text{[from eqn. (24)]}$$

$$= 0{\cdot}601\angle -28°$$

where I_1, I_2, . . . I_r, I_{r+1}, . . . are the currents through the capacitors in the 1st, 2nd, . . . rth, $(r+1)$th, . . . sections.

Also $$\frac{I_1}{I_{x0}} = 0{\cdot}469 + j0{\cdot}282 = 0{\cdot}547\angle 31°$$

Hence

$$\sum_{r=1}^{r=\infty} \hat{I}_r^2 = \hat{H}_{y0}^2\frac{\delta^2}{4}(0{\cdot}547)^2 [1 + (0{\cdot}601)^2 + (0{\cdot}601)^4 + (0{\cdot}601)^6 + \ldots]$$

$$= \hat{H}_{y0}^2\frac{\delta^2}{4}\frac{0{\cdot}300}{1 - (0{\cdot}601)^2} = 0{\cdot}469\hat{H}_{y0}^2\frac{\delta^2}{4}$$

The capacitor current I_r is equal to the total z-directed current through the rth conductor element of cross-sectional area h^2, the elements being numbered from the surface. Each current consequently experiences a resistance of $1/\sigma h^2 = 4/\sigma\delta^2$ per unit length in the z-direction. Hence the total loss at all depths of the conductor below a surface area of $1 \times \delta/2$ is given by

$$\tfrac{1}{2}\sum_{r=1}^{r=\infty} \hat{I}_r^2\frac{4}{\sigma\delta^2} = 0{\cdot}469\hat{H}_{y0}^2/2\sigma$$

Therefore the total loss per unit surface area $= 0{\cdot}469\hat{H}_{y0}^2/\sigma\delta$. This is to be compared with the analytical value of $\hat{H}_{y0}^2/2\sigma\delta$. The discrepancy of 6% in the analogue value may be reduced by the choice of a smaller interval h; thus, taking $h = \delta/4$ leads to a value of $0{\cdot}490\hat{H}_{y0}^2/\sigma\delta$ for the loss per unit surface area.

The total magnetic flux Φ per unit z-width in the conductor is obtained from eqn. (22), i.e.

$$\Phi = \frac{V_A - V_B}{K} = \frac{1}{K}\left(\frac{I_{x0}R}{2} + \frac{I_1}{j\omega C}\right)$$

$$= \frac{I_{x0}}{j\omega CK}\left(\frac{j\omega CR}{2} + \frac{I_1}{I_{x0}}\right)$$

i.e. $$\hat{\Phi} = 0{\cdot}709\hat{I}_{x0}/\omega CK$$

But $$C = \frac{h^2\sigma}{K} = \frac{\delta^2\sigma}{4K}$$

herefore $$\hat{\Phi} = 0{\cdot}709\hat{H}_{y0}\frac{\delta}{2}\frac{4}{\omega\sigma\delta^2}$$

$$= 0{\cdot}709\mu\hat{H}_{y0}\delta$$

The analytical value is $0{\cdot}707\mu\hat{H}_{y0}\delta$.

(6.2) Single Conductor Filling Rectangular Slot

The case of a single conductor, height 2δ, filling a rectang slot reduces to a 1-dimensional problem if it is assumed that magnetic flux paths in the slot are straight lines normal to sides. Fig. 6 represents a section of the conductor of width

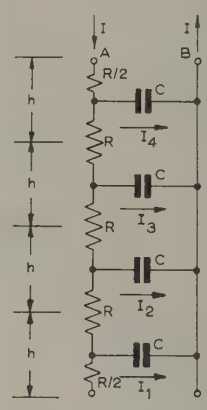


Fig. 6.—Network analogue for a single conductor filling a rectang slot, ignoring field-fringing.

and height 2δ, i.e. $h = \delta/2$. Since the magnetic field stren H_y, at the root of the slot is zero, the equivalent networ open-circuited at its lower end. Consequently the total curr I, through the capacitors, which must be equal to the cur through the corresponding section of the conductor, enter terminal A and leaves at B. To determine the effective resistance of the conductor it is necessary to find the divisio this total current between the several capacitors, as pointed in Section 5.1.

Circuit analysis shows that

$$\hat{I}_2^2/\hat{I}_1^2 = |1 + j0{\cdot}5|^2 = 1{\cdot}25$$

$$\hat{I}_3^2/\hat{I}_1^2 = |0{\cdot}75 + j1{\cdot}5|^2 = 2{\cdot}81$$

$$\hat{I}_4^2/\hat{I}_1^2 = |-0{\cdot}25 + j2{\cdot}875|^2 = 8{\cdot}33$$

Therefore $$\sum_{r=1}^{r=4} \hat{I}_r^2 = 13{\cdot}39\hat{I}_1^2$$

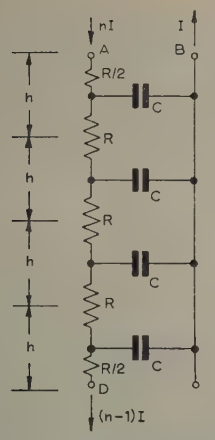
Also $$\hat{I}^2/\hat{I}_1^2 = |2{\cdot}5 - j4{\cdot}875|^2 = 30{\cdot}02$$

Hence $$\frac{R_{AC}}{R_{DC}} = \frac{4\sum_{r=1}^{r=4} \hat{I}_r^2}{\hat{I}^2} = 1{\cdot}78$$

Solution of this problem analytically without the finite differe approximation yields a value of 1·90.

(6.3) Loss in a Particular Conductor in a Rectangular Sl

This is the familiar problem of a number of identical c ductors of rectangular cross-section, each carrying a curre the conductors lying directly one above the other in a rectang slot. The ladder network of Fig. 7 is the analogue representa of the nth conductor from the root, the conductor height b taken as δ for this particular calculation and the inte $h = \delta/4$. The analogue therefore represents a section of conductor, $\delta/4$ wide by δ high, and the total capacitor curre is a corresponding fraction $\delta/(4 \times$ slot width) of the condu current i. The ratio of the magnetic field strengths at the up and lower edges of the conductor is $n/(n-1)$, since this is



ig. 7.—Network analogue for the *n*th conductor from the bottom of a slot carrying a number of rectangular conductors stacked in line.

atio of currents flowing in the slot conductors below these evels. Consequently, this must also be the ratio of the current ntering at terminal A to that leaving at terminal D. This leads o the terminal currents indicated in Fig. 7.

Circuit analysis along the lines indicated in the previous Section ives $R_{AC}/R_{DC} = 1{\cdot}076 + 0{\cdot}299(n^2 - n)$ as compared with $_{AC}/R_{DC} = 1{\cdot}084 + 0{\cdot}323(n^2 - n)$ by field analysis.

(6.4) Partially Transposed Conductors

In medium-sized machines a form of transposition depicted n Fig. 8 is often adopted. The four conductors A, B, C and

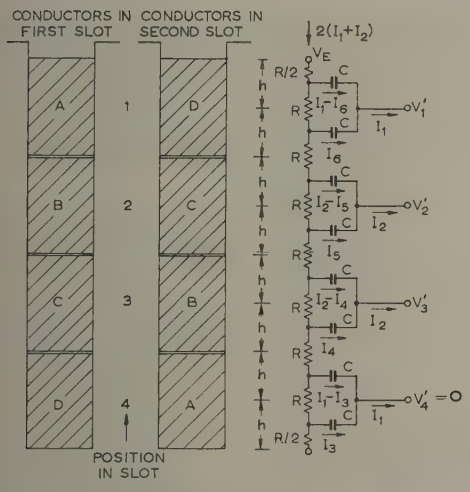


Fig. 8.—Analogue representation of four conductors partially transposed in a slot.

run through two slots between parallel connection points, the rder of the conductors in the second slot being a simple reversal that in the first. In such a case conductors A and D will rry equal currents, i_1, and conductors B and C will carry rrents i_2. The equivalent analogue has been drawn on the sis of a relatively coarse mesh, since it is the intention here ly to illustrate the setting up of the analogue. The chosen terval, h, is half the height of a conductor; hence each conductor is represented by only two network sections and the analogue currents I_1 and I_2 are a fraction (h/conductor width) of the actual conductor currents i_1 and i_2. The magnitudes and phases of the applied voltages V_E, V_1', V_2', V_3' and V_4' must be so adjusted that the indicated equalities in terminal currents are obtained. There is a further condition arising from the parallel connection. Ignoring the voltage drops in the end connections, for each slot the sum of the voltage drops per unit length in conductors A and D must equal the same sum for conductors B and C. In Section 4 it has been shown that in the analogue this must lead to

$$(V_E - V_1') + (V_E - V_4') = (V_E - V_2') + (V_E - V_3')$$

i.e.
$$V_1' + V_4' = V_2' + V_3'$$

Since voltage differences only are significant, V_4 may be set equal to zero without imposing any further condition on the analogue.

Hence
$$V_1' = V_2' + V_3', \text{ if } V_4' = 0$$

The currents I_2, I_3, I_4, I_5 and I_6 can be expressed in terms of I_1 using the above voltage relationship and the four mesh equations from the network. Such a coarse mesh does not justify eddy-current loss determinations for conductors of reasonable height, say δ, but it is interesting to note that for this height the analogue indicates that the ratio of $\hat{\imath}_2$ to $\hat{\imath}_1$ is rather less than 0·5.

To obtain significant loss data a much smaller interval would be required. Although it would still be possible to calculate the current distribution, it would be lengthy and tedious. This is one of the problems which will be solved by a practical analogue.

(7) PRACTICAL CONSIDERATIONS

To avoid undue complication in drawing the analogy between the network and the electromagnetic field, only one scaling factor K [eqn. (18)] has been introduced. K enables R and C in the analogue to be so chosen that the impedance level is satisfactory, and by assigning a dimension (m/s) to K the correspondence is made dimensionally correct. There is no reason, however, why other purely numerical scaling factors should not be introduced to permit greater convenience in operation. Two instances will be mentioned in which further scaling is either advantageous or possibly even necessary. In the first case, by operating the network analogue at a frequency higher than that of the field problem, the CR product in the analogue is correspondingly reduced; this could well lead to cheaper capacitors of greater stability. Secondly, in large machines the total current carried by the conductors in a single slot is frequently several thousand amperes; clearly a large scaling factor would be necessary here, since the foregoing analysis has taken a 1 : 1 correspondence between conductor and analogue capacitor currents. It should be noted, however, that the expression for eddy-current losses in the form R_{AC}/R_{DC} [eqn. (23)] does not depend on the magnitude of the total current but only on its distribution between the several capacitors. Since the analogue contains only linear elements the total current may be chosen arbitrarily in that case.

No consideration has been given here to such general analogue problems as the adjustment of mesh values to suit curved boundaries or a geometry involving non-integral spacing, since these features have been treated in the published work of Liebmann[6] and others.

(8) ACKNOWLEDGMENT

The author is grateful to Professor A. Tustin for his encouragement and advice throughout the study.

(9) REFERENCES

(1) FIELD, A. B.: 'Eddy Currents in Large Slot-Wound Conductors', *Transactions of the American I.E.E.*, 1905, **24**, p. 761.
(2) FIELD, M. B.: 'Idle Currents', *Journal I.E.E.*, 1906, **37**, p. 83.
(3) TAYLOR, H. W.: 'Eddy Currents in Stator Windings', *Journal I.E.E.*, 1920, **58**, p. 279.
(4) FARR, H. K., and WILSON, W. R.: 'Some Engineering Applications of the Electrolytic Field Analyser', *Transactions of the American I.E.E.*, 1951, **70**, Part II, p. 1
(5) ATKINS, W. T. J.: 'An Improved Electromagnetic Analog *Proceedings I.E.E.*, Monograph No. 263 M, Novem 1957 (**105** C, p. 151).
(6) LIEBMANN, G.: 'Resistance-Network Analogues with Une Meshes or Subdivided Meshes', *British Journal of Ap Physics*, 1954, **5**, p. 362.

DISCUSSION ON THE ABOVE MONOGRAPH

Mr. R. S. Mamak (*communicated*): The equations of eddy-current losses were first formulated by Field, and solved for the case of single-dimensional current and field. However, it can be fairly simply shown by the following example that the assumption that the field goes straight across the slot is contrary to the

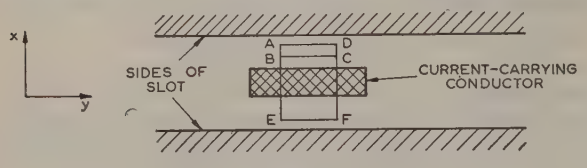


Fig. A

laws of electromagnetism. Along the path ABCD, no current is enclosed; consequently $\oint_{ABCD} H \,.\, dl = 0$. With the assumption that only H_x exists, this means that $H_{AB} = -H_{DC}$. It follows that $\oint_{AEFD} H \,.\, dl = 0$, but this is impossible, since it encloses finite current.

However, the results as obtained by Field's analysis were found to be in close agreement with practical results, and his assumption was consequently accepted. What is not so well known is that Steidinger[A] did attempt to obtain a more rigorous analytical solution by solving the two-dimensional form of the diffusion equation [eqn. (6) of the paper]. In practical cases this meant that Field's α has to be multiplied by another constant β, i.e. $R_{ac}/R_{dc} = \alpha\beta$. The practical implications of Steidinger's modification have already been fully discussed by Vallarta.[B]

As Steidinger's analysis shows, even for rather simple boundary conditions the analytical solution of the diffusion equation becomes wellnigh impossible: quantitative analysis has to be resorted to, where the main techniques are either by analogy or numerical methods, and obviously all the usual digital versus analogue-computer considerations still apply.

A powerful numerical technique for solving differential equations is the replacement of the continuous differential equation by a number of 'finite difference' equations, of the form of eqn. (9) of the paper. These equations may be collected and expressed in matrix form. It will then be seen that, essentially, it is necessary to invert a complex matrix. Mr. Roberts has inverted this matrix by using an ingenious *RC* network. However, for those not possessing an analogue computer, various powerful numerical methods of inverting matrices are available.[C] Finite-difference equations may also be obtained for all the boundary conditions discussed in this paper—in fact the finite-difference equations for a problem very similar to that illustr in Fig. 4 have been formulated, and it is hoped to publish t in the near future.

Finally, in specifying the boundary conditions for the ductors in a slot [Figs. 4(*a*) and (*b*), lines ABCD and A′B′C it is not clear what assumptions are being made with respe the current-carrying conductors in the slots on either side o one shown. If the currents in the neighbouring slots are equal and opposite to those in the slot under consideration condition that the lines of force follow the pattern illustrate Fig. 4 is incompatible with the statement that the iron is infin permeable. Consequently if the effect of the conductors in other slots is to be considered, the inherent antisymme existing in most machines should be utilized, though, of co this would entail taking a batch of slots.

REFERENCES

(A) STEIDINGER, W.: *Archiv für Elektrotechnik* 1923–24, p. 149.
(B) VALLARTA, M. S.: Discussion, *Transactions of the Amer I.E.E.*, 1924, p. 249.
(C) MILNE, W. E.: 'Numerical Solution of Differential E tions' (Wiley, 1953).

Mr. J. Roberts (*in reply*): Mr. Mamak draws attention to shortcomings of the one-dimensional analysis used by Field the conductor in an armature slot, but it is well to note tha pointed out in Section 1, this is a justifiable assumption rectangular conductors which substantially fill the width of slot and do not approach closely its open end. This condi is met in many practical designs.

Fig. 4 shows a single isolated slot with the approxi magnetic flux pattern. The flux lines AB and CD, whose dist from the mouth of the slot is equal to the length of the air- will be very nearly straight. Considering now the effe neighbouring slots carrying conductors, additional flux wil superimposed on that shown in Fig. 4(*a*). The only portio this flux which will change the current distribution in the ductors shown will be that which penetrates these conduct The effect will be small if the air-gap is small, since then additional flux will fall off rapidly with depth in the slot, a only likely to produce current redistribution in a top condu which closely approaches the mouth of the slot. To take acc of this effect on a single slot representation would involve application of suitably different potentials to the lines A′B′ C′D′ in the analogue. The current leaving the network via capacitors would not then be supplied equally from the two works representing the adjacent air-gaps.

3.013

The Institution of Electrical Engineers
Monograph No. 342
Aug. 1959
©

URFACE-INTEGRAL METHODS OF CALCULATING FORCES ON MAGNETIZED IRON PARTS

By C. J. CARPENTER, M.Sc.(Eng.).

paper was first received 19th February and in revised form 29th April, 1959. It was published as an Institution Monograph in August, 1959.)

SUMMARY

ıe force exerted on an iron part in a magnetic field may be calcu- l by a number of different methods, all of which give equivalent e distributions confined principally to a surface. Seven different ıods, giving rise to five different surface force intensities, are ⁄n to result in the same total force. Additional volume force ponents are obtained and a new expression for the actual distribu- of force, consistent with all of the methods of calculation, is ⁄ed. A common interpretation of the well-known Maxwell field ses in iron is shown to be incorrect. When the field is specified erically the calculation of force in terms of surface integrals has ıin advantages over the virtual-work method, and it is shown that le analytic expressions for force which are usually obtained in this may be easily derived as surface integrals. Practical applications ⁄onsidered.

LIST OF SYMBOLS

B = Magnetic flux density in vacuo.
B' = Magnetic flux density in iron.
H = Magnetic field strength in vacuo.
H' = Magnetic field strength in iron.
J = Conduction current density.
$M = B - \eta_0 H$ = Polarization.
η_0 = Primary magnetic constant.
μ = Permeability.
ρ = Magnetic pole density.

Suffixes

s = Surface density.
n = Component normal to surface.
t = Component tangential to surface.

(1) INTRODUCTION

ın important phenomenon underlying the operation of most tromagnetic apparatus is the mechanical force which is rted on an iron part in a magnetic field. Such forces are ally explained and calculated in terms of the virtual-work ıciple,[1, 2, 3] which provides a powerful method of analysis ling to simple expressions in terms of rates of change of uctance when saturation and hysteresis may be ignored. en this is not the case, and when the magnetizing windings distributed, the virtual-work method is less simple, and it ds to produce inaccurate results when the small differences in l energy produced by the virtual displacement have to be luated numerically. It is also open to objection in principle, e it suggests that, when the magnetization is non-linear, the e is essentially dependent on the changes produced by a dis- cement of the iron part concerned, rather than on the field figuration associated with its specified position.

ın alternative method of approach is to express the mechani- force in terms of an integral over a surface (usually the iron surface) supplemented in some instances by an additional volume integral. Perhaps the best-known example is the calculation of the force on the equivalent magnetic poles in the iron surface, but other methods (including the energy principle) may be used, as described in Section 2. Most of these methods are well known, but their practical application appears to have received little attention. One possible reason for this is that, in many problems in which the virtual-work method leads to a very simple calculation when the field configuration is suitably idealized, the way in which the same result may be obtained by surface integration is less obvious. An example is the force tending to pull two poles, or teeth, into alignment (Section 5.1). In addition, the simplicity of the surface-integration principle tends to be obscured by the fact that the different methods of calculating force in this way lead to several different expressions for the surface force intensity. The equivalence of these expressions is by no means obvious, and is confused by the fact that most, if not all, have been interpreted by different writers as representing the actual force distribution within the iron. The differences have led to a number of discussions of apparent anomalies (e.g. References 4–7), and in other instances the different methods have been incorrectly applied.

The object of the paper is to examine the different surface-integral expressions for force on iron parts and to investigate their equivalence and practical application. The methods are compared and assessed only in so far as they offer means of calculating total forces; a review of the mechanism by which the force is exerted and of the actual force distribution within the iron requires too lengthy a treatment to be included. The discussion relates specifically to iron parts, since forces on ferromagnetic materials are of the greatest practical importance, but all of the methods, and the equations deduced, may be applied by analogy to the forces acting on dielectrics in electric fields.

The M.K.S. rationalized system of units is employed. In one respect, namely the substitution of η_0 for μ_0 and μ for μ_r, the symbolism is that suggested by Cullwick and Carter.[20] This change is designed to avoid the unfortunate confusion between the properties of polarized materials and the electromagnetic relations in vacuo which has become associated with the M.K.S. system.

(2) EXPRESSIONS FOR SURFACE FORCE

(2.1) Representation of Magnetized Iron by Pole and Current Distributions

The most direct method of calculating forces on magnetized iron is to integrate the average forces on the dipoles which are the source of magnetization. This results, however, in a volume* as well as a surface distribution of force[8, 9] (whether or not the permeability is constant or the iron carries conduction currents), and it is therefore of no practical importance as a means of calculating total force. Excluding also the direct application of

ɔrrespondence on Monographs is invited for consideration with a view to ication.
r. Carpenter is at the Imperial College of Science and Technology, University of ıon.

* When curl $H = 0$ the method gives the same force distribution as the third Maxwell stress system discussed in Section 2.5.

the virtual-work method, there are at least seven other different ways in which expressions for force can be obtained.

First, any distribution of poles or currents, or both, which, when put in place of a piece of magnetized iron, reproduces the magnetic-field conditions at all points outside the iron, must experience the same total mechanical force as the part which it replaces. A simple example is provided by the method of images, in which the force acting on a large iron plate due to any system of magnetic sources is the force which would act on the image sources. A more general form of equivalent source[9, 10] is a volume distribution of poles of density

$$\rho = -\operatorname{div} \boldsymbol{M} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1a)$$

combined with a surface distribution of density

$$\rho_s = M_n \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1b)$$

This is a complete equivalent source provided that the iron carries no conduction currents and contains no 'free poles'. An alternative equivalent source[10] subject to the same restriction is a volume distribution of currents of density

$$\boldsymbol{J} = \frac{1}{\eta_0} \operatorname{curl} \boldsymbol{M} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2a)$$

combined with a surface density

$$J_s = \frac{1}{\eta_0} M_t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2b)$$

where the direction of J_s is at right angles to M_t.

To illustrate the significance of these equations consider a cylinder of iron magnetized uniformly in a direction parallel to its axis. Eqn. (1*b*) states that the polarization is equivalent to a distribution of equal but opposite magnetic polarities over the two end faces, and eqn. (2*b*) shows that the iron may also be replaced by a uniformly wound thin coil forming a current sheet coincident with the curved surface. If the polarization is not uniform and the permeability is not constant the surface distribution must be supplemented in each case by a volume distribution inside the cylinder. Both these sources are equivalent in that, when put in place of the iron, they reproduce the magnetic field conditions at all points outside the iron surface. Inside this surface (where there is now no iron) the fields cannot be the same; in the one case the magnetic field strength but not the flux density is reproduced, whereas in the other the flux density in the model, but not the field intensity, corresponds to that in the iron. The total mechanical force (if any) acting on the two equivalent sources must be the same (although the fo distributions are wholly different), and this force must also the same as that which acts on the iron.

In general, an infinite number of equivalent sources may devised in which both poles and currents are combined. T most important is that obtained by 'terminating' the field at t surface of the iron, or, in general, at any surface external to but including no other sources, so that the field is everywhe zero inside the surface chosen.[11] The equivalent source th defined consists of a surface distribution of poles of density

$$\rho_s = \eta_0 H_n \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3$$

where H_n is the normal component of the field on the outsi of the surface, combined with a surface current of density

$$J_s = H_t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3$$

directed at right angles to the tangential component, H_t, of t external field. This 'magnetic model' of the iron is not subje to the restrictions of the previous two but is a complete equiv lent of all the field sources enclosed within the surface. particular, it is immaterial whether the permeability of the ir is constant, single-valued or otherwise; in all cases the pol and currents are confined entirely to a surface.

(2.2) Forces on Equivalent Pole and Current Distributions

The surface tractions associated with each of the three equiv lent sources defined by eqns. (1)–(3) may be expressed in ter of the magnetic field intensity at the surface; the value immed ately outside the surface is chosen, since this is the same for ea of the three 'magnetic models'. The field vectors on the tw sides of the iron surface are related by the continuity condition

$$B_n' = B_n \qquad \mu H_n' = H_n \quad . \quad . \quad . \quad (4$$
$$H_t' = H_t \qquad B_t' = \mu B_t \quad . \quad . \quad . \quad (4$$

The surface pole intensity in the equivalent pole distributi defined by eqn. (1) is therefore [from eqn. (1*b*)]

$$\rho_s = \eta_0(\mu - 1)H_n' = \eta_0\left(1 - \frac{1}{\mu}\right)H_n$$

and the tangential component of the surface traction is given b

$$F_t = \rho_s H_t$$

from which the expression given in Table 1, line 1 is obtaine The normal component of the field is discontinuous, having

Table 1

	Method	Assumptions	Surface force components*		Volume force†
			Tangential, F_t	Normal, F_n	
1	Surface poles	μ constant, surface coincident with iron surface	$\eta_0\left(1 - \frac{1}{\mu}\right)H_nH_t$	$\frac{1}{2}\eta_0\left(1 - \frac{1}{\mu^2}\right)H_n^2$	$\frac{1}{\mu}\rho\boldsymbol{H}' + \eta_0(\boldsymbol{J} \times \boldsymbol{H}')$
2	Surface currents		$-\eta_0(\mu - 1)H_nH_t$	$\frac{1}{2}\eta_0(\mu^2 - 1)H_t^2$	$\mu\rho\boldsymbol{H}' + \eta_0\mu^2(\boldsymbol{J} \times \boldsymbol{H}')$
3	Combined surface poles and surface currents	None	$\eta_0H_nH_t$	$\frac{1}{2}\eta_0(H_n^2 - H_t^2)$	0
4	Virtual work	As 1 and 2	0	$\frac{1}{2}\eta_0\left(1 - \frac{1}{\mu}\right)(H_n^2 + \mu H_t^2)$	$\rho\boldsymbol{H}' + \eta_0\mu(\boldsymbol{J} \times \boldsymbol{H}')$

* Field components are those on outside of surface (in vacuo). Normal (n) direction positive outwards.
† Due to 'free polarity' ρ and conduction current $\boldsymbol{J}$.

lue H_n outside and H_n/μ inside, so that (by subtraction of that rt of the field which is due to the surface poles themselves) the rmal component of force is

$$F_n = \tfrac{1}{2}\rho_s H_n\left(1 + \frac{1}{\mu}\right)$$

om which the expression given in the 4th column of Table 1 derived.

The same reasoning may be applied to obtain the components force on the surface currents given by eqn. (2*b*), except that re the tangential component of the field is discontinuous.

In both cases the surface traction gives the entire force on e iron when there are no conduction currents in the iron and hen the permeability is uniform [so that the volume distribuons given by eqns. (1*a*) and (2*b*) are zero]. The volume force mponents which are due to non-uniform permeability may be adily derived, but they have no practical utility. The force mponents due to conduction currents in the iron are given in ction 3.

The third equivalent source consists of only a surface distribuon, as given by eqn. (3), and the force associated with this ves the whole of the force on the iron in all circumstances. he calculation of this surface traction is complicated by the ct that both tangential and normal components of the field re now discontinuous. This difficulty is most easily circumented by supposing that the magnetic poles and currents given y eqns. (3*a*) and (3*b*), respectively, lie on two different surfaces parated by a negligible distance. Only one component of the eld is then discontinuous at each surface, and the discontinuity presents the contribution made by the respective source. ach of the surface forces may then be derived in exactly the me way as before. The result obtained by adding them—hich is the same whether the poles or the currents are assumed be on the inside or the outside—is given in line 3 of Table 1.

The three surface tractions obtained in this way are entirely ifferent from each other. They will, however, give the same otal force (as is proved in Section 4), provided that the assumpons on which the first two depend are valid.

(2.3) Surface Force in Terms of Energy

The energy principle can be applied in a well-known anner[10, 12] to calculate not only total force but also an equivent force distribution. Assuming that there are no conduction urrent or 'free pole' sources included within the iron, and that xternal sources are constant, the rate of change or total potential nergy when an iron part is moved in any arbitrary direction, , is

$$\frac{\partial W}{\partial x} = \tfrac{1}{2}\eta_0 \iiint \frac{\partial \mu}{\partial x} H^2 dv$$

here the integral extends over the volume of the iron part only. When the permeability, μ, is uniform the force density correponding to the integrand is zero everywhere except at the iron rface. The surface force intensity may be calculated by ssuming that the permeability varies in a continuous manner rough the surface. The result, which is everywhere normal the surface, is given in Table 1, line 4.

(2.4) Maxwell Stresses in Vacuo

The total force on any group of field sources may be found by rawing a closed surface round them and integrating the Maxwell eld stresses in vacuo over this surface.[13] The stresses consist f a tension along the lines of force of magnitude $\frac{1}{2}\eta_0 H^2$ and n equal pressure at right angles to them. Resolving in the ormal and tangential directions relative to an arbitrarily chosen surface,[14] the component of stress directed away from the surface is $F_n = \frac{1}{2}\eta_0(H_n^2 - H_t^2)$ and the tangential component is $F_t = \eta_0 H_n H_t$.

This force is the same as that on the combined surface pole-surface current equivalent source (Table 1, line 3), so that these two methods of calculating force are essentially the same. However, in the Maxwell stress method the above components represent the force transmitted through the surface to the sources inside, whereas in the equivalent-source method the force acts on the surface. Both methods give the force on an iron part irrespective of saturation or hysteresis, and both give zero resultant force if the surface encloses no field sources.

(2.5) Second Maxwell Stress System

Maxwell[13] proposed a system of field stresses in the interior of dielectrics and these may be applied by analogy[8, 11, 15] to magnetically polarizable media (whether or not the permeability is constant). The tensile component of stress is $\frac{1}{2}\mu\eta_0(H^1)^2$ along the lines of force, again combined with a transverse pressure of the same magnitude. Thus the components resolved in the normal and tangential directions relative to any interior surface in the iron are

$$F_n = \tfrac{1}{2}\mu\eta_0[(H_n')^2 - (H_t')^2] \quad . \quad . \quad . \quad . \quad (5a)$$

$$F_t = \mu\eta_0 H_n' H_t' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5b)$$

where primes denote field values in the interior of the iron.

When the permeability is uniform and the iron contains no conduction current or 'free pole' sources, then, by direct comparison of the two stress systems, the integration of eqn. (5) over any closed surface within the iron gives zero resultant force, i.e. there is no volume force acting on the iron. It follows that the difference between the field stress on the inside of the iron surface and that on the outside gives the same result when integrated as does the exterior stress alone. Thus the force associated with the differential surface traction is equal to the force on the iron. The expression which is obtained for the surface force is the same as that given by the energy method (Table 1, line 4).

The volume force associated with the interior field stress is no longer zero when the permeability is not uniform and when conduction currents are present. Under these circumstances eqn. (5) does not give a zero result when integrated over the inside of the iron surface, and hence the differential surface traction no longer gives the force on the iron. In effect, the component of force which is subtracted when the interior stress is taken into account must be subsequently added back again as a volume force (conduction currents are taken account of in this way in Section 3.4). Clearly, for the purposes of force calculation, there is no significant practical advantage to be gained by taking the interior field stress into consideration. It eliminates the tangential component of the surface force, but this is, in any case, entirely negligible for all normal values of permeability.

(2.6) Third Maxwell Stress System

Maxwell[13] deduced a further system of 'interior' field stresses specifically in terms of magnetically polarizable media. Assuming a constant scalar permeability, the stress consists of a tensile component along the field lines of magnitude $\eta_0(\mu - \frac{1}{2})(H')^2$ combined with a transverse pressure of magnitude $\frac{1}{2}\eta_0(H')^2$. This resolves into components normal and tangential to a surface:[14]

$$F_n = \tfrac{1}{2}\eta_0[(2\mu - 1)(H_n')^2 - (H_t')^2] \quad . \quad . \quad . \quad (6a)$$

$$F_t = \mu\eta_0 H_n' H_t' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6b)$$

This stress system is mentioned in the literature less often than the previous one, but where it is described it is customarily applied to the calculation of force in exactly the same way, i.e as a surface integration only (see, for example, Reference 16). In fact, the differential surface traction, which is

$$F_n = \tfrac{1}{2}\eta_0 H_n^2\left(1 - \frac{1}{\mu}\right)^2 = \frac{1}{2\eta_0} M_n^2 \quad . \quad . \quad . \quad (7a)$$

$$F_t = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7b)$$

does not represent the total force on the iron.[8, 9, 11] There is, in addition, a force on each volume element even if the permeability is constant and there are no conduction currents. Maxwell's derivation is based specifically on a volume force whose x-component is

$$F_x = M_x\frac{\partial H_x'}{\partial x} + M_y\frac{\partial H_y'}{\partial x} + M_z\frac{\partial H_z'}{\partial x} + J_y B_z' - J_z B_y' \qquad (8)$$

in terms of the polarization, $\boldsymbol{M}$, and the conduction current, $\boldsymbol{J}$, 'free poles' being excluded.

The calculation of total force from eqn. (6) may be clarified by applying it to an interior volume element of the iron.*

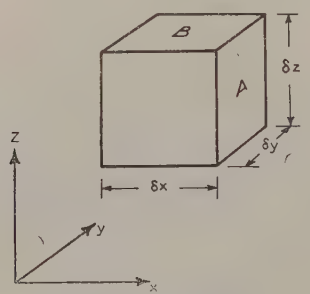


Fig. 1.—Volume element.

Taking an element of side δx, δy, δz (Fig. 1), the force in the x-direction on side A is, from eqn. (6*a*),

$$(\delta F_x)_A = \tfrac{1}{2}\eta_0\left[(2\mu - 1)\left(H_x' + \frac{\partial H_x'}{\partial x}\delta x\right)^2 - \left(H_y' + \frac{\partial H_y'}{\partial x}\delta x\right)^2 - \left(H_z' + \frac{\partial H_z'}{\partial x}\delta x\right)^2\right]\delta y\,\delta z$$

The contribution to the force on the element, in the same direction, due to face B is, from eqn. (6*b*),

$$(\delta F_x)_B = \mu\eta_0\left(H_z' + \frac{\partial H_z'}{\partial z}\delta z\right)\left(H_x' + \frac{\partial H_x}{\partial z}\delta z\right)\delta x\,\delta y$$

and the contributions from the remaining four faces are obtained in the same way. Adding the six components, the total force reduces to

$$\delta F_x = \eta_0\left[(2\mu - 1)H_x'\frac{\partial H_x'}{\partial x} - H_y'\frac{\partial H_y'}{\partial x} - H_z'\frac{\partial H_z'}{\partial x}\right]\delta x\delta y\delta z + \mu\eta_0\left[H_z'\frac{\partial H_x'}{\partial z} + H_x'\frac{\partial H_z'}{\partial z} + H_y'\frac{\partial H_x'}{\partial y} + H_x'\frac{\partial H_y'}{\partial y}\right]\delta x\delta y\delta z \qquad (9)$$

which gives a force per unit volume in accordance with eqn. (8) when curl $\boldsymbol{H}' = \boldsymbol{J}$ and div $\boldsymbol{H}' = 0$.

* This is essentially a matter of differentiation of the stress tensor, but the elementary calculation is given for clarity. It is important to distinguish this from the slight, but important, modification used in Section 4.

Thus the volume element is not in equilibrium under the in ence of the field stress only. By superposition, the force on interior part of the iron is given by integrating eqn. (6) over bounding surface, and, in particular, the total volume fo acting on the iron is given by the integral of this equation o the inside of the iron surface. The component of force whic attributed to the interface itself [eqn. (7)] is obtained by s tracting this same integral from the integrated field stress on outside of the surface. To obtain the total force the volu component and the surface component must be added toget when the two interior stress integrations cancel out and le the force associated with only the exterior field stress—whic the force on the iron.

The third Maxwell stress method is therefore useless a practical means of calculating total force. The surface fo given by eqn. (7) is not the force on the iron, and when component which must be added to it is evaluated as a surf integral the result is identical with the surface force obtai when the interior stress is ignored.

(2.7) Other Field Stresses

As Heaviside,[11] O'Rahilly[15] and others have pointed o there are many possible field stresses applicable to the inter of polarized media. Each different stress will be associat with a different volume force, but the total force acting on the ir will, inevitably, be given correctly provided only that the str reduces to the proper value in vacuo (which is unique except a constant).

For the purposes of practical calculation of total force, th 'interior' stresses are of interest only when they are associat with a zero volume force. The only stress which satisfies t condition is that described in Section 2.5, and even this gi the result as a surface integral only when there are no cc duction currents and the permeability of the iron is constant.

(3) ADDITIONAL FORCES DUE TO CONDUCTION CURREN

(3.1) Conduction Currents and Magnetic Poles

Three of the surface expressions given in Table 1 depend the assumption that the iron contains no field sources oth than the polarization dipoles, i.e. they will in general give or that component of the total force which acts on these dipol The iron will experience additional forces if, besides bei magnetized, it conducts electric current or if it has any 'fr poles' immersed in it, i.e. any magnetic polarity not account for by the polarization dipoles. The conduction currents fo a possible source of this polarity in that they may be replac by their 'magnetic-shell' equivalents, and a distributed polari as well as a distributed conduction current will therefore assumed to be present for the sake of generality. This is obvious importance in the analogous electrostatic problem.

In practice, any currents which have a significant effect the mechanical force will normally be confined to copper co ductors placed in tunnels or slots, in which case the currents m be properly regarded as outside the iron. However, it may convenient to lump together the forces acting on the copper a on the iron by ignoring the slot as such, i.e. the slot surfa In addition, the iron itself may occasionally carry large co duction currents.

In one case only, namely line 3 of Table 1, the surface fo expression is sufficient. Each of the remaining three expressio is obtained by subtracting from this result different force co ponents (obtained by integrating over the interior of the ir surface) which are associated with the additional field sourc and these components have to be evaluated as volume integr and added back again to obtain the total force.

(3.2) Volume Force in Surface-Pole Model

The required volume forces may be calculated by supposing at the current and free polarity which is distributed over each it volume of the iron is concentrated in a cavity of negligible e. In the surface-pole equivalent model of the iron it is nvenient to choose the shape of the cavity so that the total rface polarity on its walls is negligible when it contains no free larity; this implies a shape defined by two parallel surfaces a gligible distance apart and oriented so that they are also rallel to both the field and the current vector. Under these nditions there is no resultant force on the cavity walls when the rrent is placed midway between them, and the additional rce due to the current which is not accounted for by the inte- ation over the outer surface of the iron is simply the force on e (imaginary) current-carrying conductor, i.e.

$$\boldsymbol{F} = \eta_0 \boldsymbol{J} \times \boldsymbol{H}' \quad . \quad . \quad . \quad . \quad . \quad (10a)$$

r unit volume, where $\boldsymbol{J}$ is the conduction current density id the cavity field is related to the field $\boldsymbol{H}'$ in the iron by n. (4*b*).

If the iron contains free polarity of density ρ, then, when is is placed in the cavity, the additional force per unit volume similarly

$$\boldsymbol{F} = \rho \boldsymbol{H}'$$

owever, since div $\boldsymbol{H}' = \rho/\mu\eta_0$ there is induced polarity on the vity walls of the opposite sign and of magnitude

$$(1 - 1/\mu)\rho$$

that the total force acting, in addition to that accounted for y the integration over the exterior surface, is

$$\boldsymbol{F} = \frac{1}{\mu}\rho \boldsymbol{H}' \quad . \quad . \quad . \quad . \quad . \quad (10b)$$

(3.3) Volume Force in Surface-Current Model

It is most convenient to choose the cavity so that, when it ntains no conduction current, the total surface current in the alls is negligible. The cavity will therefore be supposed to ave the same shape as the previous one, but to be oriented so at the field vector is perpendicular instead of parallel to it. When a magnetic pole is introduced it induces no resultant rface current in the cavity walls, so that the additional force is e force on the pole only. From eqn. (4*a*), this is

$$\boldsymbol{F} = \mu\rho \boldsymbol{H}' \quad . \quad . \quad . \quad . \quad . \quad (11a)$$

When a conduction current is present only that component hich is normal to the field can be replaced by a concentrated rrent within the cavity; however, it is only this component hich gives rise to an additional force. The force on the quivalent conductor is $\boldsymbol{F} = \mu\eta_0 \boldsymbol{J} \times \boldsymbol{H}'$. In addition, there re currents having the same sense induced in the cavity walls nd these will add to the force. Since curl $\boldsymbol{B}' = \mu\eta_0 \boldsymbol{J}$ the agnitude of the additional current is $(\mu - 1)\boldsymbol{J}$, and the total rce additional to that resulting from the integration over the terior surface is

$$\boldsymbol{F} = \mu^2\eta_0 \boldsymbol{J} \times \boldsymbol{H}' \quad (11b)$$

(3.4) Volume Force in Energy Model

Neither eqn. (10*a*) nor eqn. (11*b*) gives actual forces on con- uction currents in the iron. They represent essentially arbitrary mponents of total force which are subtracted from the field ress in vacuo to obtain the surface traction. Similarly the pparent volume force which corresponds to the surface traction given on line 4 of Table 1 is that which is associated with the second Maxwell field stress on the interior of the iron surface. By comparison with the stress in vacuo, the volume intensity of this internal force is seen to be

$$\boldsymbol{F} = \rho \boldsymbol{H}' + \boldsymbol{J} \times \boldsymbol{B}' \quad (12)$$

and when this is not zero it must be integrated and added to the surface force.

The same result may be obtained using the energy method by considering the change in energy when the currents and poles are moved through the medium. This means that eqn. (12) gives the total mechanical force acting on field sources immersed in liquid or gaseous media. This is not, however, a purely magnetic force, but consists of the sum of a magnetic force and a hydrostatic pressure exerted by the surrounding fluid (e.g. Reference 12, p. 103). The ratio of these two components depends on the shape of the conductor in which the current flows, in the same way as the force on a conductor in a solid medium depends on the shape of the cavity made to contain it, and eqn. (12) gives no more insight than eqn. (10) or eqn. (11) into the actual magnetic forces associated with the conduction charges.

The volume force expressions are summarized in Table 1.

(4) EQUIVALENCE OF FORCE EXPRESSIONS

(4.1) Method of Verifying Equivalence

The four expressions for force listed in Table 1 give entirely different force distributions both over the surface and within the volume of the iron, but all should integrate to give the same total force in all applications where the permeability is constant. The equivalence may be verified directly by calculating, by each of the methods, the force on an arbitarily chosen volume element of iron.

When the iron carries no conduction currents and its permeability is uniform, the apparent volume force is zero in all of the methods considered except the third Maxwell stress method (whose results are not included in Table 1). But if a volume element is isolated from the surrounding iron by cutting a narrow cavity, one of whose walls forms the surface of the element, a surface distribution of force must be taken into consideration. The element becomes an independent part on which the force acting can be measured experimentally and determined theoretically by each of the methods described. The calculation is relatively simple, provided that the width of the cavity is small, so that a negligible volume of material is removed and the field conditions are not appreciably modified anywhere except in the immediate vicinity of the cavity.*

(4.2) Combined Surface-Pole/Surface-Current Equivalent

The force expression given on line 3 of Table 1 is the most general. Choosing a Cartesian volume element, as in Fig. 1, and expressing the force in terms of field values in the cavity, the component of force in the x-direction on face A is

$$(\delta F_x)_A = \tfrac{1}{2}\eta_0 \left[\left(H_x + \frac{\partial H_x}{\partial x}\delta x \right)^2 - \left(H_y + \frac{\partial H_y}{\partial x}\delta x \right)^2 - \left(H_z + \frac{\partial H_z}{\partial x}\delta x \right)^2 \right] \delta y \delta z$$

The component acting on face B is

$$(\delta F_x)_B = \eta_0 \left(H_z + \frac{\partial H_z}{\partial z}\delta z \right) \left(H_x + \frac{\partial H_x}{\partial z}\delta z \right) \delta x \delta y$$

* This also implies that the calculated force is unaffected by the cavity, and is therefore the actual force in the volume element. The method therefore appears to solve the long-standing problem[11,13] of the actual force distribution. This is, however, irrelevant in the present context and will be discussed elsewhere.

and expressions may be similarly written down for the forces on the four other faces. By adding the six components a general expression for volume force may be obtained, but since the result is required only for the purposes of comparison it is sufficient to assume constant permeability. Expressing the field components in the cavity in terms of the components in the iron [eqn. (4)] the force on the element in the x-direction becomes

$$\delta F_x = \eta_0\left(\mu^2 H_x'\frac{\partial H_x'}{\partial x} - H_y'\frac{\partial H_y'}{\partial x} - H_z'\frac{\partial H_z'}{\partial x}\right)\delta x\delta y\delta z$$

$$+ \mu\eta_0\left(H_x'\frac{\partial H_z'}{\partial z} + H_z'\frac{\partial H_x'}{\partial z} + H_x'\frac{\partial H_y'}{\partial y} + H_y'\frac{\partial H_x'}{\partial y}\right)\delta x\delta y\delta z$$

Substituting curl $\boldsymbol{H}' = \boldsymbol{J}$ and div $\boldsymbol{H}' = \rho/\mu\eta_0$, the x-component of the force per unit volume reduces to

$$F_x = \mu M_x\frac{\partial H_x'}{\partial x} + M_y\frac{\partial H_y'}{\partial x} + M_z\frac{\partial H_z'}{\partial x} + \rho H_x' + J_y B_z' - J_z B_y' \quad (13)$$

which is somewhat similar, but is not identical, to the expression deduced by Maxwell and used as the basis for his third stress system (Section 2.5)

The torque acting on the element can be similarly calculated from the tangential components of force on four of the faces. The result contains no first-order terms, so that the torque per unit volume is zero for a sufficiently small volume element.

(4.3) Alternative Methods of Calculating Force on Element

All the alternative methods of calculating force require the addition of a surface force and a force acting on the interior part of the element. The simplest is that given in line 4 of Table 1 (i.e. the virtual-work or second Maxwell stress method), in terms of which the force on face A is the x-direction is

$$(\delta F_x)_A = \tfrac{1}{2}\eta_0\left(1 - \frac{1}{\mu}\right)\left[\left(H_x + \frac{\partial H_x}{\partial x}\delta x\right)^2\right.$$

$$\left. + \mu\left(H_y + \frac{\partial H_y}{\partial x}\delta x\right)^2 + \mu\left(H_z + \frac{\partial H_z}{\partial x}\delta x\right)^2\right]\delta y\delta z$$

(the field values being those in the cavity) and the only other face which contributes is the opposite one. The force per unit volume acting on the entire surface is therefore

$$F_x = \mu M_x\frac{\partial H_x'}{\partial x} + M_y\frac{\partial H_y'}{\partial x} + M_z\frac{\partial H_z'}{\partial x}$$

in terms of field components in the iron. The appropriate volume force is given by eqn. (12). When the two components are added the total force per unit volume is in accordance with eqn. (13). The torque is clearly zero.

The force expressions given on the first two lines of Table 1 may be applied in exactly the same way, except that in both instances there will be components in the x-direction acting on all six faces and in the interior of the element. By either method the resultant is in accordance with eqn. (13).

All the methods considered are therefore equivalent for a small volume element, and it follows that they must be equivalent for an iron part consisting of any number of such elements. The surface force components will cancel out at mutual interfaces and leave a resultant surface distribution over the outer surface of the iron only, plus a volume force distribution given by eqn. (10), (11) or (12), respectively.

(5) APPLICATION OF SURFACE-FORCE EXPRESSIONS

(5.1) Choice of Surface of Integration

The first Maxwell stress, or combined surface-pole and surface-current, method has obvious advantages over the other forms of surface integral listed in Table 1, and this will almost invariably, the most suitable method for prac problems. One of its principal advantages does not appea be sufficiently widely utilized, namely that the surface of inte tion need not be chosen to coincide with the surface of the part on which the force is required. Many examples of f calculation which give simple analytic results when appropr assumptions are made are customarily analysed by the en method, and yet the same result may be obtained as simply surface integration, provided that the surface is suitably cho

A simple example is the force reaction between two sal poles, or teeth, tending to pull them into alignment (Fig. Assuming that the poles have infinite permeability and that t is a uniform-field region between adjacent pole tips and ligible field intensity at the remote tips, the alignment force

$$F = \tfrac{1}{2}\eta_0 H^2 g \quad . \quad . \quad . \quad . \quad .$$

where H is the field strength in the uniform-field region and the gap length. This result is obtained by the energy met by Moullin[4] and Tustin.[17] One method of verifying it a surface integral is to choose a surface such as ABCD [Fig. 2

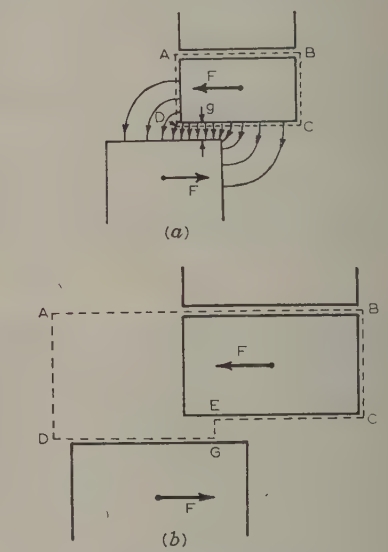


Fig. 2.—Force on overlapping poles.
(*a*) Usual integration surface.
(*b*) Modified integration surface.

which coincides with the pole faces on three sides, and is clos by supposing an indefinitely small air-gap within which plane AB may be drawn. This fictitious gap is necessa since the force given by eqn. (14) cannot be measured exp mentally without including the other iron parts to which pole is attached, or by making an actual air-gap. Under assumptions specified there is no resultant component of s ways force contributed by the parts DC and CB of the surfa and there is no resultant force contribution due to AB, provi that the 'fringe' field of the main gap is negligible at the point The sideways force is therefore wholly accounted for by face AD, and for this reason the force is customarily supposed act on the side of the pole (or tooth).

When the surface is chosen in this way the sideways forc difficult to calculate analytically (although not numerical But the difficulty disappears for other surfaces, both in this in similar problems. If, for example, the plane AD is separa

m the pole surface by a distance sufficient to bring it outside fringe field, then it no longer contributes to the force, which w acts on the DC plane. The resultant integral, which can evaluated by parts, gives a force in accordance with eqn. (14). owever, it is more convenient to localize the force to one par- ular part of DC, and this can be done by drawing the surface own in Fig. 2(*b*). Here DC is interrupted at GE so that DG incides with the surface of the lower pole and EC with that the upper one. DG and EC then contribute nothing to the ce, which is therefore given by the normal pressure on the rt GE; this can be evaluated by inspection and is in accordance th eqn. (14). To simplify the result further it is desirable to arate BC from the pole surface as well as AD, so that B lies a negligible field. There is then no resultant force on AB due the fringe field of the lesser gap.

The surface-integration method has been found to be applic- le in an equally simple manner to every other example which s been examined and in which an analytic result is obtainable the virtual-work method. A further illustration of the choice surface for a transverse-force, or torque, calculation is shown Fig. 3. The device illustrated, which is described by Tustin,[17]

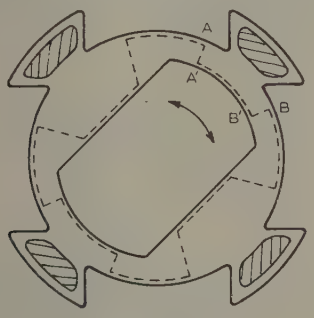


Fig. 3.—Torque motor.

nsists of a stator with two windings in slots at right angles to ch other and magnetizing a salient-pole rotor. The rotor is ojected to a torque which may be expressed as a function of e field intensities at AA′ and BB′, provided that the fringe lds set up at the various circumferential discontinuities do not teract, and provided that the permeability is sufficiently large. he same result follows by inspection from the surface of inte- ation shown by the broken lines.

(5.2) Numeric Integration of Surface Force

When simplifying assumptions such as those made in deriving qn. (14) are not justified, the energy method, as customarily pplied,[1–3] requires a numeric field solution before and after the rtual displacement. The displacement made must be large ough to produce a significant change in either flux or mag- etomotive force. The change in field energy is then calculated; general this involves integrating the flux/current curve for each rn of the magnetizing winding, and then summing the results r each of the field plots. When, on the other hand, the force is btained as a surface integral, only one field solution is required* nd no flux-current integration is necessary. Moreover, the umeric calculation may consist of only a minor modification f an analytic solution. For example, if the flux density on BC Fig. 2 is not negligible, the additional force component due to it ay be very simply and rapidly estimated and subtracted from qn. (14) once the actual field distribution is known.

In practice, the information from which the force is obtained will usually be given in the form of a field plot, as, for example, when the electrolytic tank is used. The surface integral must then be evaluated in terms of measured angles and distances, and the simplest finite-difference form to use is that suggested by Lehmann.[18] Suppose one part of the surface chosen is an equipotential such as PQ (Fig. 4), which might coincide with the

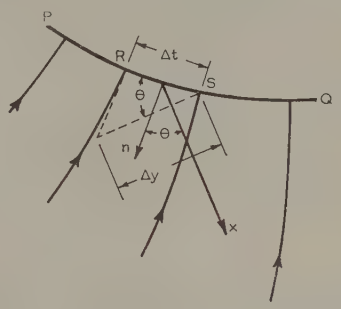


Fig. 4.—Numerical evaluation of surface integral.

surface of the iron when the permeability is large. Let the flux tube, RS, make an intercept Δt with the surface and represent a flux $\Delta\Phi$. From Table 1, line 3, the component of force on RS resolved in any arbitrary direction x at angle θ to the normal to RS, is

$$\Delta F_x = \frac{1}{2\eta_0}\left(\frac{\Delta\Phi}{\Delta t}\right)^2 \Delta t \cos\theta$$

Hence the total force on PQ in this direction is

$$F_x = \frac{1}{2\eta_0} \sum_{PQ} \frac{(\Delta\Phi)^2}{\Delta y} \quad . \quad . \quad . \quad . \quad . \quad (15a)$$

where Δy is drawn at right angles to the x-direction and

$$\Delta t = \Delta y \cos\theta \quad . \quad . \quad . \quad . \quad . \quad (15b)$$

The flux increment, $\Delta\Phi$, is usually a constant and may be taken outside the summation, so that the field analysis reduces to a series of measurements of Δy. Exactly the same construction may be used if PQ is a flux line instead of any equipotential, and RS represents an increment ΔF_m in m.m.f. The surface force is again normal to PQ and has the same magnitude, but its direction is reversed. Hence

$$F_x = -\tfrac{1}{2}\eta_0 \sum_{PQ} \frac{(\Delta F_m)^2}{\Delta y} \quad . \quad . \quad . \quad . \quad (16)$$

In problems in which an attempt is made to allow for saturation the magnetic equipotentials may have an appreciable inclination relative to the iron surface. Where this occurs along parts such as DG and EC in Fig. 2(*b*) of the surface of integration, it is most convenient to modify the latter so as to follow a zigzag path along the field equipotentials and flux lines, as in Fig. 5. If the potential increment along PP′ is such as to divide the flux tubes into curvilinear squares, the force acting on the whole of PQ may be obtained by summing eqn. (15*a*) [or

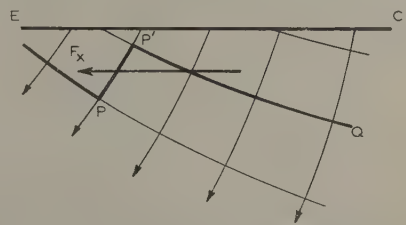


Fig. 5.—Integration surface near saturated iron.

* This may be no direct advantage when calculating a force/displacement curve. here is, however, an advantage in accuracy, since no differencing is necessary.

eqn. (16)] from P to Q with constant $\Delta\Phi$. The Δy-increments are interpreted (with appropriate signs) in the same way for both PP′ and P′Q.

(5.3) Torque on a Slotted Rotor

In the combined surface-pole and surface-current equivalent model the current density is given by eqn. (3*b*), i.e. $J_s = H_t$. Thus the line integral of H_t round any closed contour drawn in the surface gives the total surface current linking this contour. But, by Ampère's law, the line integral of the field intensity is determined by the conduction current flowing in the iron, from which it follows that the equivalent current, J_s, corresponds to the actual current in the iron, but distributed over the surface in accordance with the above equation.

It follows that the torque on an ordinary slotted machine rotor may be expressed in terms of the torque which would act on the rotor conductors if they were placed in the air-gap. The rotor is magnetically equivalent to a distribution of poles and currents over a cylindrical surface passing through the tops of the teeth, the intensity of these sources being defined by eqn. (3). When the permeability of the iron is infinite, each slot pitch in this surface contains a total current equal in magnitude to the corresponding slot current. The torque on the rotor is the same as that on this current, provided that the magnetic polarity is regarded as placed on the inside of the cylindrical surface and the current on the outside. When the permeability is not infinite the surface current differs slightly from the slot current.

The principle of equivalent surface distributions thus provides an exact basis for the orthodox treatment of slotted machines, in which the mean torque is shown to be the same as that which would act on the rotor currents when placed on the surface of an equivalent cylindrical rotor. The more exact model shows how these surface currents have to be distributed, in the actual air-gap field, so as to produce the same torque in all rotor positions as well as the same mean value. One result which follows from this is that the torque pulsations corresponding to the slot pitch do not disappear when the slots are closed, even if the permeability of the iron remains sufficiently large to be regarded as infinite. The effect of the slots is negligible only when they are sufficiently far from the surface for the tangential field component at the surface to be uniform over each slot pitch.

(5.4) Forces acting on Toroidal Electromagnet

Five different surface-force expressions are obtained when eqn. (7) is added to those listed in Table 1, and of these four differ by only very small amounts (i.e. by factors of $1/\mu$ or less) when the field at the surface of the iron has no tangential component. Because these differences are small, apparent anomalies arise if the attractive force tending to close an air-gap is assumed to act on the iron surfaces forming the two sides of the gap. Several writers[4, 5, 7, 16, 19, 20, 21] either analyse examples of such forces incorrectly or else point out apparent anomalies, the example generally treated being the split toroidal electromagnet (Fig. 6). It is therefore instructive to calculate the force tending to pull the two halves of such a magnet together, using each of the expressions listed in Table 1 and taking account of all terms, whether small in magnitude or not. The calculations are simple if the core is assumed to be rectangular in section and if the air-gap is sufficiently short compared with its area for the fringe field to be negligible (i.e. the calculation refers to the limiting case as the gap is closed). Since the flux density is uniform it is reasonable to assume constant permeability.

The force is most easily obtained from the equivalent surface poles, since, for the toroidal configuration, the surface polarity is confined to the two sides of the air-gap. Hence, from Table 1,

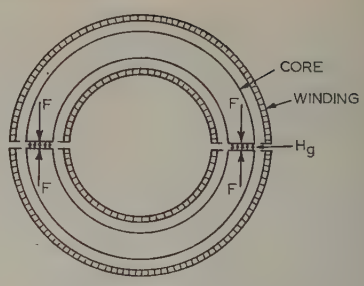


Fig. 6.—Toroidal electromagnet.

line 1, the restraint which has to be applied to the iron at e air-gap is, per unit cross-section,

$$F_1 = \tfrac{1}{2}\eta_0\left(1 - \frac{1}{\mu^2}\right)H_g^2 \quad . \quad . \quad . \quad . \quad (1$$

where H_g is the field in the gap. If the magnetizing windin similarly split into two halves these are also subjected t force tending to bring them together. This force is very ea calculated (by the Maxwell stress method) since it is unaffec by the presence of the core; it is

$$F_2 = \tfrac{1}{2}\eta_0\frac{1}{\mu^2}H_g^2 \quad . \quad . \quad . \quad . \quad . \quad (1$$

at each gap, per unit cross-section. Hence the combined fo on both the core and winding is

$$F_1 + F_2 = \tfrac{1}{2}\eta_0 H_g^2 \quad . \quad . \quad . \quad . \quad . \quad (1$$

in accordance with the Maxwell stress in the gap.

In terms of the equivalent surface currents the force on iron is confined to the cylindrical surface. The force on t equivalent coil may be written down by inspection by the sa method used to obtain eqn. (17*b*), or it can be obtained applying the appropriate surface-force expression in Table taking moments about the end of the half-core and integrati to find the force applied at the other end. By both methods t force is in accordance with eqn. (17*a*).

In terms of the combined surface-pole and surface-curre equivalent the force given by eqn. (17*c*) acts on the end fa of the iron. This is not, however, the force acting on t half-core, as there is an additional contribution from the tw curved surfaces. The resultant of these is outwards, since t field strength is greater on the inside surface, and its value given by eqn. (17*b*), as can be shown by taking moments a integrating, or by inspection from the transverse Maxwell stre or other considerations. The total force on the iron is ther fore in accordance with eqn. (17*a*).

According to the virtual-work method, the force which ac on the end-face is

$$\tfrac{1}{2}\eta_0\left(1 - \frac{1}{\mu}\right)H_g^2$$

Again this must be supplemented by a radial force on the curv surfaces, giving in this case an inwards resultant and therefo increasing the pull. The contribution which this makes at t gap may be calculated by taking moments, as before, and t same total force is obtained. The example also shows o advantage to be gained by applying the energy method in t form of an equivalent surface force; when it is applied direc by considering the effect of a virtual displacement of the ha core[21] the energy changes at the two curved surfaces are diffic to allow for and are easily overlooked.

The problem is analysed by du Bois,[19] Howe[5] and Carter[20] the third Maxwell stress system, and the force on the iron culated by applying eqn (7) to the end-face only. The result not the actual force on the iron, but does correspond to that mponent of it which is due to the other half-core, as distinct m the effect of the magnetizing winding. However, the ysical significance of this result is fortuitous and the difference tween eqns. (7) and (17*a*) is explained entirely by the fact that e former does not represent the force on the iron. The axwell stress methods are, in principle, unable to distinguish tween components of force due to different external field urces.

(5.5) Forces Acting on Straight Electromagnet

In the toroidal electromagnet the force acting on each half-re is accounted for by the force on the equivalent poles in the r-gap surfaces. In general this will be true only when there is component of flux density normal to the iron surface at any her point in the system, and it is of interest to examine the plication of the different surface-force expressions to an ectromagnet in which this condition is not satisfied. A mple example consists of an infinitely long iron rod cut into o by an air-gap which is again supposed to have a negligible ngth/area ratio (Fig. 7). The rod is magnetized by a uniform

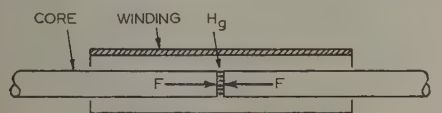


Fig. 7.—Straight electromagnet.

inding whose length is finite, but which is sufficient to establish uniform field in the region of the air-gap. For this arrange-ent the force acting on each half-core is given most simply by e integral of the surface force whose tangential component is ro, so that there is no contribution from the cylindrical surface hat is by the virtual-work method). From Table 1, the force on e iron, per unit cross-section, is

$$F = \tfrac{1}{2}\eta_0\left(1 - \frac{1}{\mu}\right)H_g^2 \quad . \quad . \quad . \quad . \quad (18)$$

where H_g is the field in the gap); this is slightly different from e force which acts on the core of the toroidal electromagnet qn. (17*a*)].

In terms of the surface-pole equivalent, the difference between ns. (17*a*) and (18) is due to the force on the magnetic polarity stributed along the cylindrical surface. This, as shown in the ppendix, can be evaluated quite simply by integration if the on is assumed to have a small cross-section, and it gives an twards component which, when subtracted from the force cting on the end-face, produces the total force given above. he remaining two surface-force expressions in Table 1 may be nilarly integrated, with the same result.

(6) CONCLUSIONS

Surface-integral methods of calculating forces, although well own in principle, have not been as widely applied in practice their merits appear to warrant. The force on an iron part n always be obtained from a single field plot by integrating er a surface, irrespective of saturation, hysteresis and the esence or absence of conduction currents. The method is rticularly simple to apply numerically; in addition, it gives alytic results as readily as the virtual-work method, provided at the surface is appropriately chosen.

The three Maxwell stress methods and four others examined lead to five different surface tractions. These have been proved to be equivalent, provided that, in all but one, an appropriate volume force distribution is added to the surface force when the iron carries conduction currents. The magnitude of the volume force differs in the four cases. Only that force distribution which is associated entirely with the surface has any practical significance as a method of calculating total force.

The system of field stresses proposed by Maxwell for the interior of iron parts is the only method which leads to volume distribution of force when there are no conduction currents present and the permeability is uniform. Because of this the surface traction associated with these field stresses does not, as is sometimes assumed, give the force on the iron. When the volume force is taken into account in a practical calculation the method reduces to the evaluation of the field stresses in vacuo only.

It is probable that none of the equivalent distributions examined corresponds to the way in which the force is actually exerted on magnetized iron. An alternative expression relating to a measurable force density, consistent with all of the methods of analysis, has been derived.

(7) REFERENCES

(1) Doherty, R. V., and Park, R. H.: 'Mechanical Force Between Electric Circuits', *Transactions of the American I.E.E.*, 1926, **45**, p. 240.
(2) Karapetoff, V.: 'Mechanical Forces Between Electric Currents and Saturated Magnetic Fields', *ibid.*, 1927, **46**, p. 563.
(3) Roters, H. C.: 'Electromagnetic Devices' (Wiley, 1951).
(4) Moullin, E. B.: 'Principles of Electromagnetism' (Clarendon Press, 1932), p. 180.
(5) Howe, G. W. O.: 'Stress in Electric and Magnetic Fields', *Wireless Engineer*, 1946, 23, p. 319.
(6) Larmor, J.: 'On the Theory of Electrodynamics as Affected by the Nature of the Mechanical Stresses in Excited Dielectrics', *Proceedings of the Royal Society*, 1892, **52**, p. 55.
(7) Lee, T. H.: 'Forces and Stresses in an Electromagnetic Field', *Transactions of the American I.E.E.*, 1957, 76, Part I, p. 267.
(8) Larmor, J.: 'A Dynamical Theory of the Electric and Luminiferous Medium—IV', *Philosophical Transactions of the Royal Society*, 1897, 190A, p. 205.
(9) Livens, G. H.: 'Theory of Electricity' (Cambridge University Press, 1918).
(10) Stratton, J. A.: 'Electromagnetic Theory' (McGraw-Hill, 1941).
(11) Heaviside, O.: 'Electrical Papers' (Macmillan, 1892), Articles 23, 52.
(12) Abraham, M., and Becker, R.: 'Classical Theory of Electricity and Magnetism' (Blackie and Son, 1947).
(13) Maxwell, J. C.: 'Treatise on Electricity and Magnetism' (Clarendon Press, 1904), Vol. I, Ch. 5, and Vol. II, Ch. 11.
(14) Searle, G. F. C.: 'On the Magnetic Field due to a Current in a Wire Placed Parallel to the Axis of a Cylinder of Iron', *Electrician*, 1898, **40**, p. 453.
(15) O'Rahilly, A.: 'Electromagnetics' (Longmans, Green and Co., 1938), Ch. 2, and Ch. 4, Section 5.
(16) Hague, B.: 'Electromagnetic Problems in Electrical Engineering' (Oxford University Press, 1929).
(17) Tustin, A.: 'Direct Current Machines for Control Systems' (E. and F. N. Spon Ltd., 1952), Ch. 10.
(18) Lehmann, Th.: 'The Calculation of Magnetic Attraction', *Transactions of the American I.E.E.*, 1926, 45, p. 383.
(19 du Bois, H.: 'The Magnetic Circuit' (Longmans, Green and Co., 1896), p. 155.
(20) Carter, G. W.: 'The Electromagnetic Field in its Engineering Aspects' (Longman's Green and Co.), 1957, p. 205.
(21) Moore, A. D.: 'Fundamentals of Electrical Design' (McGraw-Hill, 1927), p. 13.

(8) APPENDIX

(8.1) Calculation of Force on Cylindrical Iron Core

The force tending to pull the two halves of the core together, in the electromagnet shown in Fig. 7, may be written down by inspection from the virtual-work expression given on line 4 of Table 1. Eqn. (18) results. The three other surface-force expressions listed in the Table include tangential components which must be integrated along the cylindrical surface of the core, as discussed in Section 5.5. The details of the calculation of force on the surface pole distribution are given here to illustrate the method; the expressions listed on lines 2 and 3 of Table 1 integrate in exactly the same way.

Consider a short length δx of the core, distance x from the end of the magnetizing winding. Let the field intensity on the axis at one end of the element be H'; then at the other it is $H' + \dfrac{\partial H'}{\partial x}\delta x$. But div H' is zero, so that the radial field com-

ponent at the curved surface, when the radius of this surface is small, is

$$H_r' = \frac{\partial H'}{\partial x}\delta x\frac{\pi r^2}{2\pi r\delta x} = \tfrac{1}{2}r\frac{\partial H'}{\partial x}$$

inside the iron. The force on the magnetic polarity in the curved surface is therefore, from Table 1, line 1, and from eqn. (4),

$$\eta_0\left(1-\frac{1}{\mu}\right)H'\tfrac{1}{2}r\mu\frac{\partial H'}{\partial x}2\pi r\delta x$$

Hence the total force associated with this polarity is

$$F = \eta_0(\mu-1)\pi r^2\int_0^{H_c'} H'dH' = \tfrac{1}{2}\eta_0\pi r^2(\mu-1)(H_c')^2$$

where $H_c' = H_g/\mu$ is the (uniform) field inside the coil. This force acts outward, so that it tends to separate the two half-cores. Thus the force

$$F_1 = \frac{\eta_0}{2}\left(1-\frac{1}{\mu^2}\right)H_g^2$$

which acts on the surface polarity at the air-gap (Table 1, lin must be reduced by an amount

$$F_2 = \frac{\eta_0}{2\mu^2}(\mu-1)H_g^2$$

and the total force inwards on the core, per unit area, is

$$F_1 - F_2 = \frac{\eta_0}{2}\left(1-\frac{1}{\mu}\right)H_g^2$$

in accordance with eqn. (18).

The method of calculation shows, moreover, that the f is a function of the air-gap field only, and is independent of field distribution along the axis. Thus in this respect also result is consistent with the previous one.

DISCUSSION ON THE ABOVE MONOGRAPH

Mr. P. Hammond (*communicated*): Forces on iron parts are often very difficult to calculate, and the author has rendered a great service to designers, whose attention should be drawn particularly to the fact that the surface of integration need not be the same as the actual surface of the iron. Careful choice of the surface of integration may shorten the calculation considerably. This is well illustrated by Figs. 2 and 3 of the paper.

The author stresses the advantages of using equivalent surface distributions which reduce the field within the iron to zero and thus abolish the volume forces. These equivalent distributions are those used by Green* in his analysis of the electric field, and are often called Green's equivalent strata.

While the mathematical part of the paper is admirable, I am less happy about the physical content. Apparently the author considers that none of the methods he lists give insight into the actual force distribution. All the methods are apparently merely mathematical devices and the choice between them is a matter of convenience. Without wishing to appear ungrateful I would suggest that few engineers will be satisfied to leave the matter there. The designer must know where the forces act on the actual parts of his machine. Equivalent layers and Maxwell stresses just will not do. Nor is this request for knowledge unreasonable. There must be a definite answer in every case, and one could use strain gauges to discover it. There ought, therefore, to be a mathematical method which is physically more correct than the others.

This method, I suggest, is the method of surface and volume poles given by eqns. (1*a*) and (1*b*). It is, of course, true that dipole interactions are insufficient by themselves to account for ferromagnetic behaviour. Nevertheless on a large-scale view, which averages the effect of many magnetic domains, the magnetic forces are correctly described by a pole distribution. From energy considerations it is the free magnetism at the ends of a domain that matters. It is surely there that the forces act. On this point of view the volume distribution merely becomes another surface distribution of poles. A divergence of magnetization within the iron draws attention to a lump of sla the iron. On the surface of this lump there will be sur polarity. If we use good-quality material there will not be m divergence within the iron and so the forces will act on the surface. Although the method of surface polarity may always be the most convenient one in calculations, I w suggest that it is the method that gives the closest physical insi

Mr. C. J. Carpenter (*in reply*): Mr. Hammond has raise most interesting topic in his reference to force distribution, it is one which I should have liked to elaborate. Howeve requires a much more detailed treatment than was possibl the monograph, and I must confine my remarks to two po I do, as Mr. Hammond suggests, regard all the method analysis listed in Section 2 as mathematical devices, and I a that it is not satisfactory to leave the matter there. But paper does go further, and the force density expression der in Section 4 [eqn.(13)] does provide a possible answer.

There are several reasons for supposing that the equiva poles and the forces associated with them have no phys significance, not the least of them being that the pole i is a purely mathematical concept. The basic element polarizable material is the dipole, and the forces which ac the dipoles are not the same as those which act on the equiva distribution of poles. For example, every dipole of the mat —whether viewed at a macroscopic or at a microscopic lev will experience a force unless it happens to be situated uniform field. The equivalent pole distribution, on the c hand, is associated with only a surface force (assuming cons permeability elsewhere and no conduction currents). The face poles may be visualized as the ends of highly idealized hypothetical dipole chains, and in this view the forces whic on the poles correspond to forces on elements which consi *half* dipoles. Such an element is, of course, wholly fictitio

Mr. Hammond suggests that energy considerations sup the hypothesis of polar forces. However, the use of en considerations to obtain force distributions (as distinct from forces) is open to serious objections. In any case, as is sh in Table 1, it gives a different result from the surface pole met

* Green, G.: 'An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism' (Nottingham, 1828), Article 7.

621.385.1.029.6

The Institution of Electrical Engineers
Monograph No. 343
Sept. 1959
©

MICROWAVE TUBES—AN INTRODUCTORY REVIEW WITH BIBLIOGRAPHY

By A. F. HARVEY, D.Phil., B.Sc.(Eng.), Member.

(*The paper was first received* 24*th April*, 1958, *and in revised form* 10*th April*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in September*, 1959.)

SUMMARY

The paper reviews the various types of electron vacuum tubes employed for amplification and generation at microwave frequencies. Emphasis is placed on principles of operation and on tubes recently developed to give high power output, oscillations at the highest frequencies and low noise factors. The treatment is restricted to conventional tubes in which the output energy is derived from the d.c. input. The subject is interpreted in terms of published work, the text being closely associated with a bibliography which is complete to the Microwave Valve Convention of May, 1958.

After a general introduction, the first part discusses grid-controlled tubes. It is then shown how the interaction of space-charge waves with resonant cavities and slow-wave circuits results in the various forms of drift-space and growing-wave tubes. The second part deals with crossed-field interaction in planar and circular geometry and includes an examination of the magnetron. An account is given of novel methods of generation of submillimetre waves and the usual sources of electrons are described. The third part analyses noise phenomena in oscillators and amplifiers.

LIST OF SYMBOLS

A_0 = Direct beam current density (sheet), A/m.
B = Susceptance, mhos.
B = Magnetic flux density, Wb/m^2
c = Speed of light *in vacuo* = 3×10^8 m/sec.
C = Pierce's gain parameter.
C = Capacitance of circuit, farads.
C_{cg} = Cathode-grid capacitance, farads.
d_{gc} = Grid-cathode spacing, m.
e = Charge on electron = $1{\cdot}6027 \times 10^{-19}$ coulomb.
E = Electric field, volts/m.
E_z = Axial component of E.
f = Subscript denoting fast space-charge wave.
f = Frequency, c/s.
f_b = Bandwidth, c/s.
F = Signal/noise ratio.
g_m = Mutual conductance, mhos.
G = Conductance, mhos.
G_r = Resonant value of G.
h = Planck's constant = $6{\cdot}6252 \times 10^{-34}$ joule-sec.
H = Magnetic field intensity, AT/m.
I = Current, amp.
I_0 = Direct beam current, amp.
I_1 = Peak value of r.f. beam current, amp.
$j = \sqrt{-1}$.
J_c = Cathode current density, amp/m^2.
J_n = Bessel function of the first kind and nth order.
J_1 = Average value of r.f. beam current density, A/m^2.
k = Boltzmann's constant = $1{\cdot}38042 \times 10^{-23}$ joule/deg K.
l = Length, m.
L = Inductance, henrys.
L_c = Inductance of cathode lead, henrys.
m = Mass of electron = $9{\cdot}1066 \times 10^{-31}$kg.
m_s = Voltage gain of amplifier stage.
m_{sr} = Resonant value of m_s.
n = Refractive index.
$\mathscr{P}$ = Perveance of electron beam = $I_0/V_0^{3/2}$.
P = Power, watts.
P_a = Anode dissipation, watts.
P_e = Power given up by electrons, watts.
P_0 = D.C. input power = I_0V_0.
P_{out} = R.F. output power, watts.
Q = Space-charge parameter.
Q_l = Loaded Q-factor of circuit.
Q_u = Unloaded Q-factor of circuit.
r = Radial co-ordinate, m.
r_a = Radius of anode, m.
r_c = Radius of cathode, m.
s = Subscript denoting slow space-charge wave.
s = Reduction factor for plasma frequency.
S = Poynting vector, watts/m^2.
t = Time, sec.
T = Absolute temperature, deg K.
T_c = Cathode temperature, deg K.
T_0 = Room temperature $\simeq$ 290° K.
v = Linear velocity, m/sec.
v_0 = Average electron velocity, m/sec.
v_p = Phase velocity along circuit, m/sec.
v_1 = Peak value of r.f. electron velocity, m/sec.
V = Volume, m^3.
V = Voltage, volts.
V_a = Average anode voltage.
V_H = Hartree threshold voltage.
V_{in} = Peak value of r.f. input voltage.
V_k = Peak value of kinetic voltage of beam.
V_0 = Accelerating voltage of beam.
V_{out} = Peak value of r.f. output voltage.
V_1 = Peak value of r.f. voltage across resonator gap.
W = Energy, joules.
X = Reactance, ohms.
Y = Admittance, mhos.
Y_e = R.F. admittance of electron beam, mhos.
z = Axial distance, m.
Z = Impedance, ohms.
Z_c = Coupling impedance of circuit, ohms.
Z_e = Characteristic impedance of beam, ohms.
Z_L = Impedance of load, ohms.
Z_t = Cold impedance of tube, ohms.
Z_w = Intrinsic impedance of free space = 377 ohms.
α = Attenuation coefficient.
β = Phase-change coefficient = $2\pi/\lambda$.
β_c = Phase-change coefficient of circuit wave.
β_e = Phase-change coefficient of space-charge wave.
β_g = Gap coupling factor.
γ = Propagation coefficient = $\alpha + j\beta$.
γ_e = Propagation coefficient of space-charge wave.
ϵ = Dielectric constant.
ϵ_0 = Permittivity of free space = $(1/36\pi)10^{-9}$ farad/m.

This is an 'integrating paper'. Members are invited to submit papers in this category, giving the full perspective of the developments leading to the present practice in a particular part of one of the branches of electrical science. Correspondence on Monographs is invited for consideration with a view to publication.

Dr. Harvey is at the Royal Radar Establishment.

η = Overall efficiency.
η_a = Anode efficiency.
η_c = Circuit efficiency = P_{out}/P_e.
η_e = Electronic efficiency = P_e/P_0.
θ = Angular co-ordinate.
θ_d = Drift angle of klystron.
κ = Relativistic factor.
λ = Free-space wavelength, m.
λ_g = Guide wavelength, m.
λ_p = Plasma wavelength = $2v_0/\omega_p$.
μ_0 = Permeability of free space = $4\pi \times 10^{-7}$ henry/m.
ρ = Volume charge density, coulombs/m^3.
ρ_0 = Average value of ρ.
ω = Angular frequency, rad/sec.
ω_r = Resonant value of ω.
ω_p = Effective plasma angular frequency, rad/sec.
Rationalized M.K.S. units are used unless otherwise indicated.

Part 1. GRID-CONTROLLED, SPACE-CHARGE WAVE AND TRAVELLING-WAVE TUBES

(1) GENERAL PRINCIPLES

(1.1) Operating Parameters

Electron vacuum tubes for use at microwave frequencies include conventional as well as special types[138] peculiar to such frequencies. Amplification may be at high level in order, for example, to increase the power transmitted, or at low level to enable small signals to be detected, in which case low-noise characteristics are desirable. In either case adequate gain over the required frequency bandwidth is necessary. Such amplifiers can be made to generate power as self-oscillators by feedback from the output to the input circuit with the correct phase and amplitude.

Grid-controlled tubes operating at microwave frequencies are subject to three main effects. The first is due to inter-electrode capacitance: if this is tuned with an inductance, the output is given by

$$V_{out} = \frac{g_m V_{in}}{[(1/R) + j(\omega C - 1/\omega L)]} \quad . \quad . \quad . \quad (1)$$

The gain is

$$m_s = |V_{out}/V_{in}| = \frac{g_m R}{[1 + R^2(\omega C - 1/\omega L)^2]^{1/2}} \quad . \quad (2)$$

being 3 dB down at frequencies such that

$$R(\omega C - 1/\omega L) = \pm 1 \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Hence the frequency bandwidth is given by

$$f_b = \frac{1}{2\pi CR} = \frac{g_m}{2\pi C m_{sr}} \quad . \quad . \quad . \quad . \quad (4)$$

The product $m_{sr} \times f_b$ is termed the figure of merit and may be applied to many types of tube. For the 6AK5, a typical miniature pentode, it has the value 4×10^8 c/s.

The second effect is due to the cathode-lead inductance and the cathode-grid capacitance combining to produce an input conductance, $\omega^2 g_m L_c C_{cg}$, which, for the 6AK5 at 200 Mc/s, is about 600 micromhos. This effect can be reduced by incorporating[284] a suitable inductance in the screen lead or by using separate cathode leads for input and output circuits. The third effect is due to electron transit-time which, under static conditions, is given by

$$T = 6{\cdot}7(d_{gc}/J_c)^{1/3} \times 10^{-8} \text{ sec} \quad . \quad . \quad . \quad (5)$$

For a spacing of 0·05 cm and cathode loading of 1 A/cm^2 the transit time is $2{\cdot}5 \times 10^{-10}$ sec, representing about one-quarter period at a frequency of 900 Mc/s. If the transit time betv cathode and grid is an appreciable fraction of the r.f. per power can be absorbed in the grid circuit even when no elec current flows. The input conductance due to this powe proportional to $g_m T^2 f^2$, and for the 6AK5 tube it is al 180 micromhos at 200 Mc/s.

The frequency of oscillation of microwave generators is sub to sources of instability. The variation due to ambient temp ture of a copper resonant circuit at a frequency of 3 Gc/ about 1 Mc/s for a 20° C change. Since the frequency governed by the ratio of electronic conductance to the effec capacitance of the resonant circuit, it varies with electr voltage—the so-called frequency pushing. The frequency va tion due to the load can be studied by means of the equiva circuit in Fig. 1(*a*). The resonant circuit associated with

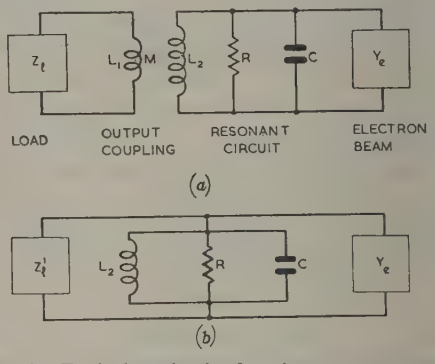


Fig. 1.—Equivalent circuit of a microwave generator.
(*a*) Equivalent circuit.
(*b*) Simplified equivalent circuit.

tube is represented by L_2, C and R, and the coupling to load by L_1 and M. This circuit is simplified in Fig. 1(*b*) assuming an ideal transformer of ratio $1 : L_2/M$ with an ind tance $L_1(1 - M/L_1L_2)$ in series with its primary and an ind tance L_2 in parallel with its secondary winding. The tra formed load impedance is given by

$$Z_L' = \left(\frac{L_2}{M}\right)^2 \left[Z_L + j\omega L_1 \left(1 - \frac{M^2}{L_1 L_2}\right)\right] \quad . \quad .$$

At frequencies near resonance the electronic admittance becor

$$Y_e = \frac{M/L_2}{Z_L + j\omega L_1\left(1 - \frac{M^2}{L_1 L_2}\right)} + \left(\frac{C}{L_2}\right)^{1/2}\left(\frac{1}{Q_u} + 2j\delta\right)$$

where $\delta = (\omega - \omega_r)/\omega_r$. From eqn. (7), putting

$$X_1 = j\omega L_1\left(1 - \frac{M_2}{L_1 L_2}\right) \quad . \quad . \quad . \quad . \quad .$$

$$Y_e = \left(\frac{M}{L_2}\right)^2\left(\frac{1}{Z_L + X_1} + \frac{1}{Z_t - X_1}\right) \quad . \quad .$$

When the tube is not oscillating, Y_e is zero and the propert of the resonant circuit and the coupling to the load may determined[132] by measurement of Z_t, the cold impedance.

At microwave frequencies, the output power and frequency an electron-tube generator will be measured in terms of amplitudes and phases of the reflection coefficient of the lo and a Rieke or circle-diagram presentation is often employ A typical diagram is shown in Fig. 2(*a*), from which it may visualized that a load of constant v.s.w.r. but variable ph

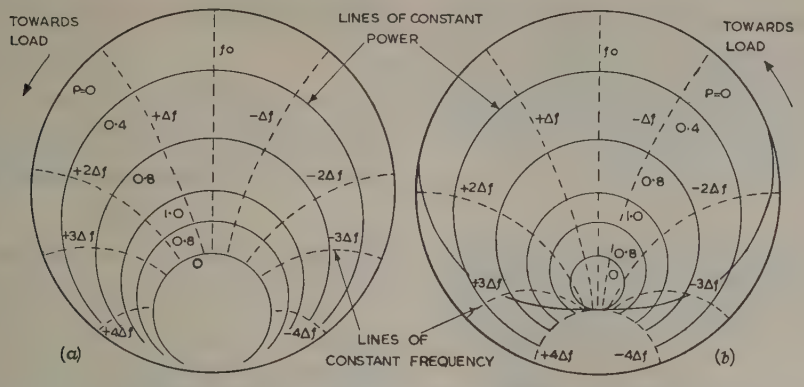


Fig. 2.—Typical power/frequency diagrams.
The variation in power, P, is shown in terms of maximum power and frequency deviation, Δf, from the mean frequency, f_0.
(*a*) Load at the generator.
(*b*) Load at 10 wavelengths from the generator.

uld cause a cyclic variation of frequency which is small near upper part of the diagram but much more rapid near the er part or 'sink'. The effect of a long transmission line can included by rotating every point on the diagram through the propriate clockwise or counterclockwise angle. Fig. 2(*b*), for line of ten wavelengths, shows that the conditions near the k are considerably modified. For some load reflection efficients no oscillations are possible, whilst at others, frequency mps or splitting may occur. Such long-line effects are undeable and may be eliminated by the use of non-reciprocal rite devices.[64]

(1.2) Grid-Controlled Tubes

Grid-controlled tubes, usable at the lower microwave frencies, are usually triodes and tetrodes. The addition of a ppressor grid to form a pentode has not been found advanwith a bandwidth of 10 Mc/s. In general, tetrodes are favoured[150] below 1 Gc/s and triodes at higher frequencies.

The electrodes of grid-controlled tubes[333] may be cylindrical or planar[90, 243] and, in order to reduce the inductance, are often brought out of the vacuum envelope via disc seals. In the 'lighthouse' triode 2C40, suitable[204] for operation up to 3 Gc/s, the electrode diameters are 0·18 in, while the grid-cathode and grid-anode spacings are 0·004 in and 0·012 in respectively. Another tube[240] has a cathode area of 0·2 cm^2, cathode-grid distance (cold) of 20 μ, a grid system composed of parallel tungsten wires, diameter 7 μ, pitch 25 μ and an anode diameter of 4·5 mm. This tube gave a gain of 17 dB at 4 Gc/s with a bandwidth of 110 Mc/s and provided, as an oscillator, 150 mW at 6 Gc/s. A triode for frequencies up to 10 Gc/s is adaptable[12] for use in either lumped-constant or distributed circuits.

Grid-controlled tubes are usually associated with coaxial-

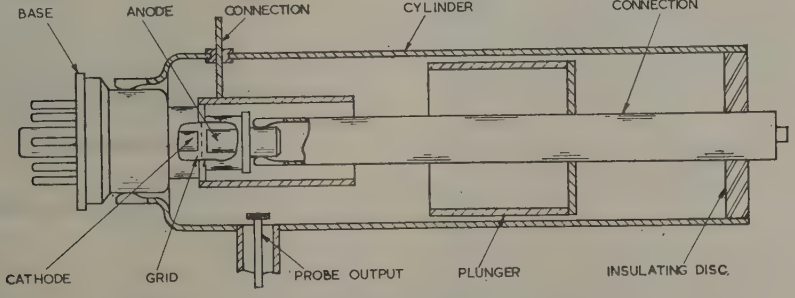


Fig. 3.—2 Gc/s lighthouse triode with re-entrant cavity.
The cylinders have spring contact fingers.

geous, since the r.f. output voltage is relatively low when the be is adjusted for optimum efficiency. High-frequency operaon is facilitated by the adoption of improved circuits such as tributed amplifiers.[32] Triodes,[5] whether acting as low- or gh-level amplifiers, are generally used in an earthed-grid rcuit[105, 161, 239, 337] in which feedback is only via the small ode-cathode capacitance. For example, a miniature 6AJ4 ode gave,[149] in such a circuit, a gain of 6 dB at 900 Mc/s line circuits.[13, 14, 69, 324, 325] Several arrangements have been developed[127, 260], and, as an example, a re-entrant triode circuit[132] is shown in Fig. 3. Radial-line and waveguide cavities are employed at the higher frequencies, and in one design[285] an annular cavity was simultaneously excited by a number of tubes placed symmetrically about the axis. An output of 500 watts at 1 Gc/s was obtained using fourteen 2C39 triodes. A small planar triode possesses[216, 246] a transconductance of 0·05 mho,

giving a gain of 10 dB at 1 Gc/s. Other low-level amplifiers have been based on pencil triodes[136] and other tubes.[54, 113, 250, 291]

The output power in the medium and high ranges[96, 155] is limited primarily by dimensions, which in turn are related to frequency, by anode efficiency and by permissible power dissipation at the anode. For the largest dimensions consistent with satisfactory operation the output is given by

$$P_{out} = \frac{0{\cdot}015 f^2 P_a \eta_a}{1 - \eta_a} \quad . \quad . \quad . \quad . \quad (10)$$

where f is in Gc/s. For example, a tetrode for 0·9 Gc/s would deliver 12½ kW with a gain of 20 and 40% efficiency with broad-band operation. A triode for the same frequency, with 50% efficiency, cathode loading 1 A/cm^2, anode dissipation 500 watts/cm^2, would have a power output of 90 kW with a bandwidth of 8 Mc/s.

A beam power tetrode for frequencies near 500 Mc/s contained[18, 19] a centrally located cylindrical anode surrounded by an array of 40 units, each containing a control grid, screen grid and cathode. With such an assembly, an interfering TE_{11} mode distribution of electric field could arise and give reduced output: this effect was overcome by balancing with external sectoral trimmer-plates. The class-B operating conditions were 8·25 kV 8 amp anode input, 900 watts grid input, giving a useful output of 27 kW at 41% efficiency. A modified tube, with lengthened envelope, gave 100 kW at an efficiency of 70% with an anode voltage of 16 kV. Developments[35, 82, 92, 183] of medium-power triodes have included an example tunable over the range 3·8–4·2 Gc/s with a gain of 4, an output power of 10 watts and a bandwidth of 100 Mc/s. Triodes have been designed[322] for use as stable oscillators.

(2) BEAM AND CAVITY INTERACTION

(2.1) Space-Charge Waves

It was first shown by Tonks and Langmuir[331] that an electron plasma or beam can support space-charge waves of an electromechanical nature. Their energy exists in two forms: electrostatic energy, which is capacitive in nature and associated with the bunching together of the electrons, and kinetic energy, which is inductive in nature and associated with differences in the velocities of the electrons. The theory of such waves shows[131, 213, 287, 309] that both TE and TM waves can propagate but only the latter produce[350] axial bunching of the electrons. If a space-charge wave of small amplitude is superimposed on an electron beam, then

$$I = I_0 + I_1 \varepsilon^{-\gamma z + j\omega t} \quad . \quad . \quad . \quad . \quad (11)$$

$$v = v_0 + v_1 \varepsilon^{-\gamma z + j\omega t} \quad . \quad . \quad . \quad . \quad (12)$$

where

$$(j\omega - v_0\gamma)^2 = e\rho_0/m\epsilon_0 \quad . \quad . \quad . \quad . \quad (13)$$

In the absence of external fields, the real part of γ is small and thus I and v vary sinusoidally with time and distance along the beam. Possible values of γ, obtained by solution of eqn. (13), are

$$\gamma = j(\omega \pm \omega_p)/v_0 \quad . \quad . \quad . \quad . \quad . \quad (14)$$

where

$$\omega_p^2 = e\rho_0/m\epsilon_0 \quad . \quad . \quad . \quad . \quad . \quad (15)$$

In a finite beam surrounded[253] by conductors, the plasma frequency is reduced by a factor s which has been calculated[36] for various geometries.

Eqn. (14) represents two waves travelling with phase velocities $v_0/(1 \pm \omega_p/\omega)$, one slightly slower and the other slightly faster than the electron beam. The group velocities of the two waves are the same and equal to the beam velocity.

The electron motion may be replaced, so far as energ concerned, by the kinetic voltage

$$V_k = -mvv_0/e \quad . \quad . \quad . \quad . \quad .$$

The real power flow in the beam is given[212] by

$$P = \tfrac{1}{2}(I_1 V_k^* + I_1^* V_k) \quad . \quad . \quad . \quad .$$

The ratio of V_k to I_1 represents a characteristic impedance. S

$$\tfrac{1}{2}mv_0^2 = eV_0 \quad . \quad . \quad . \quad . \quad . \quad .$$

the impedances of the fast and slow waves become respecti

$$Z_{ef} = V_{kf}/I_{1f} = +2\omega_p V_0/\omega I_0 \quad . \quad .$$

$$Z_{es} = V_{ks}/I_{1s} = -2\omega_p V_0/\omega I_0 \quad . \quad .$$

The electron beam behaves[29] like a transmission line and usual impedance transformation relations apply.

Since Z_{es} is negative, the power flow[212] along the bea negative, the group velocity is positive and thus the st energy[345, 346] must be negative. This is because the ch density is greatest in regions of less than average velocity least in regions of higher velocity. The opposite condit obtain with the fast wave and thus the power flow and st energy are positive. Practical electron beams are axially s metric and confined by means of a longitudinal magnetic f The conditions for the propagation of space-charge waves then modified.[192, 290]

(2.2) Amplifier Klystrons

Interaction between an electron beam and a resonant ca forms the basis of the klystron amplifier. Much of the orig analysis of this tube followed a ballistic treatment[7, 106, 107, 353 but greater accuracy can be achieved with space-charge v theory. Although radial-beam tubes with annular resona have been described,[209] the klystron usually has the a arrangement shown in Fig. 4. The input cavity sets up fast

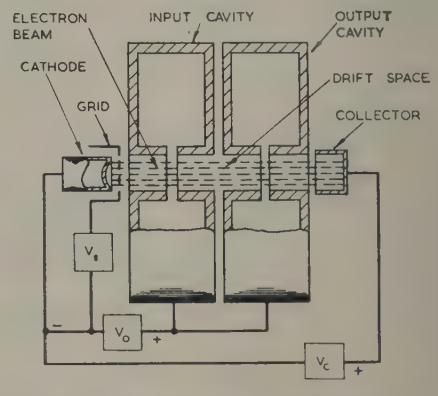


Fig. 4.—Typical klystron amplifier tube.
The electron velocity entering the buncher is $v_0 = (2eV_0/m)^{1/2}$.

slow waves, the current modulations being equal in magni and opposite in phase, thus cancelling, while the velocity mod tions are equal and in phase, thus adding. If $V_1 \ll V_0$ is r.f. voltage across the resonator gap the initial electron velo becomes

$$v = v_0\left(1 + \frac{\beta_g V_1}{2V_0}\varepsilon^{j\omega t}\right) \quad . \quad . \quad . \quad .$$

The velocity-modulated beam then enters a field-free drift s

ong which the fast and slow waves alter in relative phase, using[133] a current modulation of

$$I = I_0 + \frac{jI_1\omega}{\omega_p}\varepsilon^{-j\omega z/v_0+j\omega t}\sin\frac{\omega_p z}{v_0} \quad . \quad . \quad (22)$$

he beam bunches and debunches with distance as shown in g. 5. The first (and largest) maximum of current is seen from n. (22) to occur at a distance of $\lambda_p/4$.

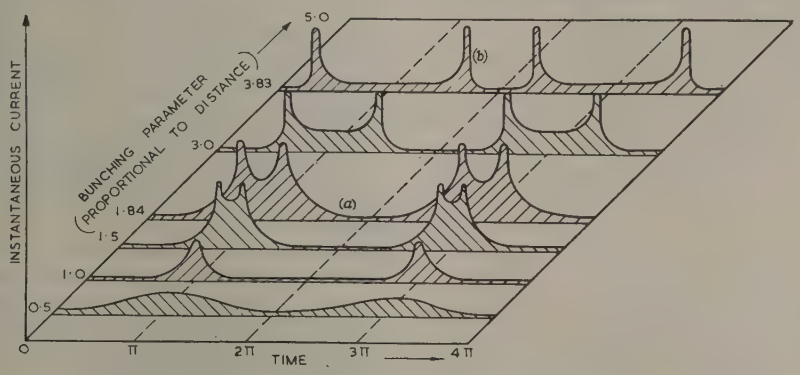


Fig. 5.—Electron bunching due to velocity modulation.
(*a*) is contour for highest efficiency.
(*b*) is contour for zero fundamental.

If the output cavity is placed at this position, its field excites nother pair of space-charge waves, the phases of the components eing such that the two fast waves cancel and the two slow waves dd. Thus, the beam leaving the second cavity possesses nega- ve energy due, of course, to the positive energy abstracted as seful output. The r.f. beam current was shown by Webster[364] o consist of an infinite number of sinusoidal alternating com- onents which are all harmonically related to the frequency pplied to the buncher. The fundamental component has an mplitude given by

$$I_1 = 2I_0\mathrm{J}_1\left[\frac{\beta_g V_1\omega}{2V_0\omega_p}\sin\frac{\omega_p z}{v_0}\right] \quad . \quad . \quad . \quad . \quad (23)$$

he term inside the brackets is the bunching parameter and for w-density beams has a value of 1·84 at the first current maxi- um. With optimum bunching the maximum efficiency is 3·4%. With a bunching parameter of 3·83 the fundamental utput is zero. These results are modified in the presence of pace charge[11, 100, 357] and a focusing magnetic field.[57]

Under large-signal conditions the r.f. output gap voltage is early equal to V_0 and it is usual, for design purposes, to compute e motion numerically for a number of electrons with different ntrance phases. An intermediate cavity may be placed along e drift space to give a cascade amplifier. The current induced this cavity by the bunched beam produces an r.f. field which, turn, produces additional velocity modulation. The increased ngth necessitates improved focusing of the beam but the gain increased and the theoretical efficiency is 80%. Klystrons ith three or more cavities can, by suitable stagger-tuning of dividual resonators, be given a broad frequency response. ne attempt to attain high efficiency involved[124] the use of oupled harmonic resonators to give non-sinusoidal bunching.

Advances in the design of klystron amplifiers have recently en reported.[206, 320, 330, 338] Practical two-cavity tubes have en described,[220] including an example[39] for a frequency of Gc/s. Multi-cavity klystrons have given[65, 205] an output of kW at 10 Gc/s and a five-cavity tube for this frequency has been developed.[55] A four-cavity tube operating at 28 kV gave[191] an output of 25 kW at 2·9 Gc/s. With synchronous tuning of the cavities the gain was 60 dB; by suitable stagger-tuning the gain was 29 dB over a bandwidth of 51 Mc/s. At a frequency of 1 Gc/s, a three-cavity klystron gave[286] 5 kW with an efficiency of 30–40% and a gain of 20–30 dB.

Very high pulse powers have been achieved with klystron amplifiers, the operating voltages being such that relativistic effects[10] must be considered in the detail design. A three-cavity tube, with a beam input of 370 kV, 190 amp gave[58] an output of 30 MW with an efficiency of 43%: the frequency was 2·857 Gc/s with a tuning range of 100 Mc/s, while the pulse length was 1 microsec and the heater power 800 watts. Similar tubes for the 10 Gc/s and 1 Gc/s bands have given peak outputs of 1–3 MW, average powers of 1–8 kW and gains of 30–40 dB: mechanical tuning over a range of 15% was possible.

(2.3) Reflex Klystrons

Amplifier klystrons can be made to oscillate by feeding back a fraction of the output power to the input circuit and adjusting the electron transit-angle to be

$$\theta_d = (4n - 1)\pi/2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

where n is an integer. More usually, the input and output interaction-gaps are combined in a single cavity to give the reflex klystron shown in Fig. 6(*a*). The electron beam, after the first transit, passes into a retarding region where a suitably shaped repeller or reflector[234] causes it to reverse and pass through the cavity a second time. Maximum power is trans- ferred from the returning bunched beam when the drift angle, adjusted by means of the reflector voltage, is such that the r.f. electric field at the gap has its maximum retarding value. Thus, as may be seen from Fig. 6(*b*), the optimum drift-angle is $2\pi(n + \frac{3}{4})$, giving rise to various modes of oscillation.[132, 264]

The power transferred from the electron beam to the resonant cavity is given[10, 11] by

$$P_e = \beta_g I_0 V_1 \sin\theta_d \mathrm{J}_1(\beta V_1\theta_d/2V_0) \quad . \quad . \quad . \quad (25)$$

for low space-charge density. For fixed values of θ_d and V_0 the power output will have a maximum given by

$$P_{max} = 0{\cdot}398 I_0 V_0/(n + \tfrac{3}{4}) \quad . \quad . \quad . \quad (26)$$

Since the beam input is I_0V_0 the electronic efficiency[207, 312] becomes

$$\eta_e = 0{\cdot}398/(n + \tfrac{3}{4}) \quad . \quad . \quad . \quad . \quad (27)$$

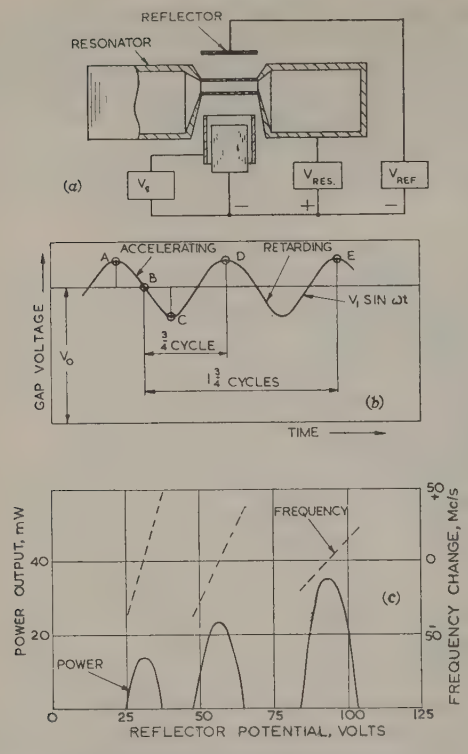


Fig. 6.—Operation of reflex klystron.
(*a*) Typical cross-section of tube.
(*b*) Optimum drift times.
(*c*) Variation of frequency and power with reflector voltage.

Eqn. (26) shows that the highest power output is obtained with a small value of n, i.e. when the reflector is at its greatest negative potential. This is illustrated in Fig. 6(*c*). The effect of space charge[60] on the performance of reflex klystrons is complicated by the presence of the returning beam. Information on the general operating conditions of these tubes has been given by a number of workers.[41, 42, 114, 137, 377] Variation of the reflector voltage from its optimum value causes the beam admittance to become complex and hence changes[26] the frequency of oscillation. This so-called electronic tuning is shown in Fig. 6(*c*) and is seen to be accompanied by variations in power output.

A number of practical reflex klystrons have been developed[140, 215, 258, 261, 288] and measurements on their properties have included bunching efficiency,[16, 17] hysteresis effects[233] and pulse operation.[77, 332] For frequencies of 2–4 Gc/s the resonant cavity is usually external[33] to the tube proper. In the 2K28 and CV35 tubes[257] the inner portion of the resonator is brought out of the glass envelope by disc seals. This technique permits a wide operating frequency and such plug-in tubes have been developed[259] for frequencies up to 10 Gc/s.

The manufacture of reflex klystrons for millimetre wavelengths[196, 197, 218] presents difficulties because of the small size of the component parts. The resonant cavity is usually placed inside the vacuum envelope, the output being taken via a sealed window to the external waveguide. Two tubes, which together covered the range 16–43 Gc/s with an output of 5–10 mW, employed[154] coaxial-line resonators. In a reflex klystron for 35 Gc/s the construction included[374] a fundamental-mode cavity. With an input of 2 kV, 10 mA, the output power was 200 mW. Low-voltage operation can be achieved by placing grids at the resonator gap and examples for 50–60 Gc/s have been described.[289] Very high frequencies can be obtained[20] by using a resonator tuned to a high harmonic of the bunching frequency.

In most reflex klystrons, mechanical tuning is accomplished by flexing the diaphragm forming one of the walls of the cavity. In, for example, the 2K25 this motion is via a screw-operated strut, but in another tube, the 2K45, it is achieved by thermal means[264] in which the strut, forming the anode of an auxiliary triode, is varied in temperature by controlling the current arriving on it. By means of a bimetallic combination the strut deforms the resonant cavity to change the frequency of oscillation. Klystrons provided with external cavities can be tuned by plungers or, alternatively, by utilizing the variable permeability of magnetized ferrites.[48] For example, in a CV2164 tube, the resonant cavity was extended[116] to accommodate a length of waveguide containing a ferrite phase changer. The reflector voltage was maintained at the optimum value for power output and tuning was achieved over 9·2–9·8 Gc/s with a power ranging between 5–25 mW.

A modification of the reflex klystron is the multiple-reflector tube[221] in which several electron transits are used to give increased efficiency. It is, of course, important that the successive bunches have the correct phase, and an efficiency of 2 is usual for frequencies up to 10 Gc/s. A typical multi-reflector klystron for 3·5 Gc/s would give 10 watts output, the range of mechanical tuning being 300 Mc/s. Another tube, which is suitable for millimetre wavelengths, consists[15, 56] basically of a single-cavity two-gap klystron with a floating drift tube. The theory of operation is similar to that of the reflex klystron but by applying a separate voltage to the drift tube, thus varying the transit angle, it is possible to produce frequency tuning. In the monotron, the drift tube is omitted and bunching takes place during the long transit across the cavity.

In the retarding-field oscillator[53, 143] the bunching takes place in a single region containing both an r.f. and a d.c. field. The reflector, as shown in Fig. 7(*a*), forms part of the resonant cavity: such a design permits simple construction and wide mechanical-tuning range. When the electrons enter this combined field the operations of velocity modulation, bunching and delivery of r.f. power to the circuit are all carried out in the same interaction space. In a tube[51, 52] for 30 Gc/s, the resonator had a diameter of 0·14 in and a height of 0·05 in. The coaxial line was 0·102 in long with inside and outside diameters of 0·065 in and 0·072 in respectively. A pointed probe, at reflector potential, was used in the reflector to aid in the focusing of the beam returning to the anode. The electron gun gave a beam of diameter of 0·01 in with a perveance of $3{\cdot}5 \times 10^{-6}$ amp/volt$^{3/2}$. The output performance of this tube is shown in Fig. 7(*b*), from which it will be seen that the power output was above 40 mW over the frequency range 27–33 Gc/s. Other tubes have given powers of 500 mW, while 300 mW was obtained at 50 Gc/s.

(3) BEAM AND CIRCUIT INTERACTION

(3.1) Growing-Wave Tubes

Modification of, or interaction between, space-charge waves in electron beams can result in their increasing in amplitude. In the resistive-wall amplifier[23] the electron beam is surrounded by a resistive film which may be in the form of either strip or hollow tubes. In the example of Fig. 8, both the input and output couplings are via resonant cavities. The microwave signal in the input cavity produces slow and fast waves in the electron beam. As each wave travels, it induces currents in the resistive wall surrounding it, with consequent dissipation of energy and decrease of power. Since the fast wave has positive power this means that its amplitude decreases to a very small

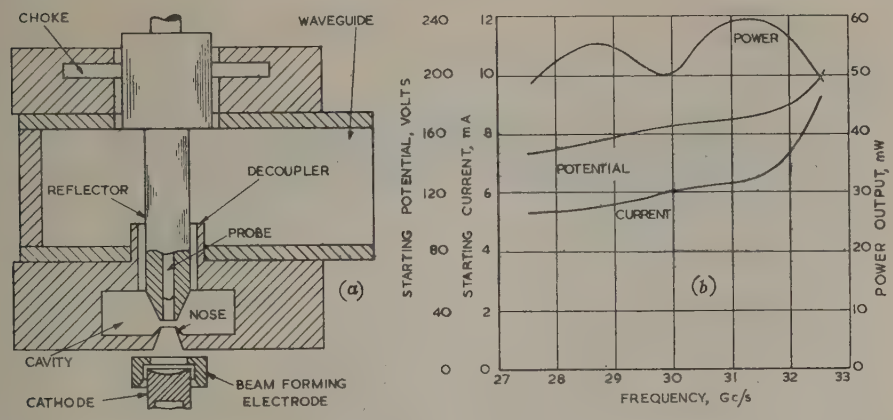


Fig. 7.—Retarding-field oscillator.

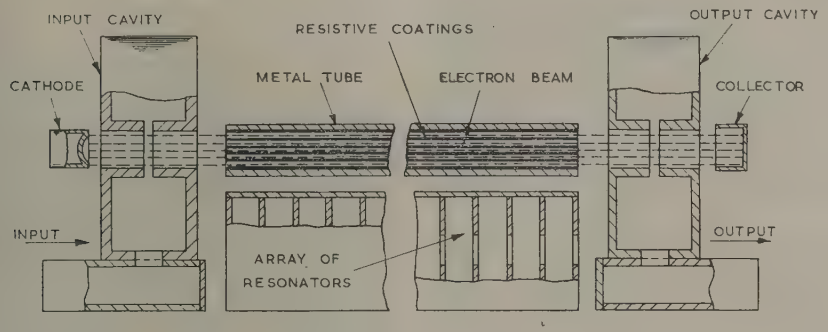


Fig. 8.—Growing-wave tubes.
Example shown is a resistive-wall amplifier with (inset) alternative 'easitron' structure.

value. The slow wave, however, has negative power and thus its amplitude increases.

In one tube the resistive layer was a coating of tin oxide about 10^{-5} in thick, supported on a thin glass tube. For a beam current of 10 mA the resistive-wall tube had a gain of 15 dB compared with the 4 dB of the same tube fitted with a plain metallic drift tube. It has been found that the capacitive-shunting effect of the glass support tends to reduce the gain. This can be partly overcome by etching the glass to a thickness of 0·005 in or by mounting[22] the resistive coating on an inductive wall.

In the 'easitron' tube[189, 272, 276] the electron beam is in close contact with a periodic structure. In the inset of Fig. 8, the series of resonators is adjusted so that, at the operating frequency, a lossless negative susceptance is presented to the beam. The effective plasma frequency becomes imaginary and the phase-change coefficients of the fast and slow waves are, on substitution in eqn. (14), found to be equal and given by

$$\beta_{ef} = \beta_{es} = (\omega + j\omega_p)/v_0 \quad . \quad . \quad . \quad . \quad (28)$$

The respective characteristic impedances are

$$Z_{ef} = \frac{+2j\omega_p V_0}{\omega I_0} = +jZ_e \quad . \quad . \quad . \quad (29)$$

$$Z_{es} = \frac{-2j\omega_p V_0}{\omega I_0} = -jZ_e \quad . \quad . \quad . \quad (30)$$

The power flow along the beam, given by

$$P = -jI_{1f}^* I_{1s} Z_e + (-jI_{1f}^* I_{1s} Z_e)^* \quad . \quad . \quad . \quad (31)$$

takes place only when both waves are present. The two waves vary in amplitude with distance, the one increasing and the other decreasing, so that $I_{1f}^* I_{1s}$ is constant. Hence, by suitable input and output coupling, the increasing wave can give amplification of microwave signals.

Methods of amplification have been proposed[151, 262] in which the characteristic impedance is varied periodically along the beam. This results[84] in a stop band when the lengths are such that

$$l_A = \frac{\pi v_0}{2\omega_{pA}} = \frac{\lambda_{pA}}{4} \quad . \quad . \quad . \quad . \quad . \quad (32)$$

$$l_B = \frac{\pi v_0}{2\omega_{pB}} = \frac{\lambda_{pB}}{4} \quad . \quad . \quad . \quad . \quad (33)$$

the subscripts A and B denoting the parameters of the alternate sections. These stop bands are characterized[276] by growing and decaying waves. Such space-charge wave amplification can also be analysed[10, 169] by considering the effect of the sudden steps on the velocity or current modulation of the beam. Eqns. (19) and (20) show that the beam impedance can be changed by variation of the accelerating voltage or of the plasma frequency by change of current density or of the factor s.

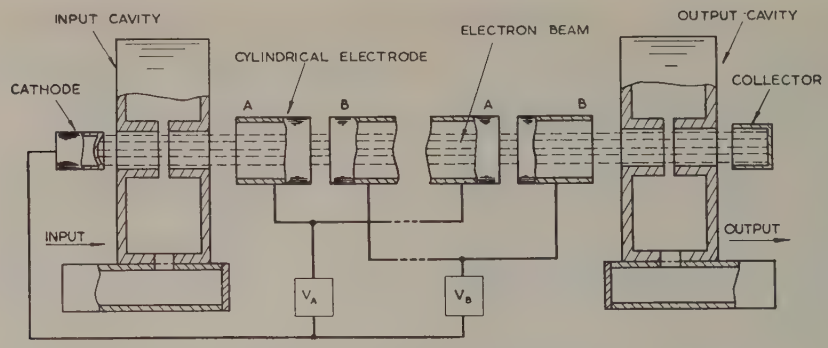


Fig. 9.—Space-charge-wave amplifier.
The regions of different potentials result in growing and decaying waves.

Amplification by change of beam potential may be achieved by the velocity-jump tube[91, 102] shown in Fig. 9. The electron beam is surrounded by a series of metallic cylinders of the correct length and at alternate steady potentials. At an upward jump, the fluctuating electron velocity is increased by the square root of the voltage ratio. The downward jump occurs at a plane of zero velocity modulation and hence has no effect. Amplification is thus achieved as the beam progresses along the structure. One tube, with helices for input and output coupling, gave a gain of 20 dB at 2·89 Gc/s with beam potentials of 66 and 680 volts.

In the stepped or rippled-wall amplifier,[22] the beam impedance is varied by alteration of the radius of the drift tube. If the tube is expanded at the plane of a current maximum, the square root of the current is increased by a factor $(s_A/s_B)^2$. Six quarter-wave sections were employed in a tube giving[24] gains in the range 19–25 dB at frequencies of 2·3–1·8 Gc/s.

The radius of the electron beam is varied in the scalloped-beam tube.[223, 224] Small-signal theory[301] shows that it is possible to increase the r.f. current density in such a beam by arranging to have a low d.c. density in the region where bunching occurs and a high density where debunching takes place. Experimental tubes have amplified at frequencies of 1–4 Gc/s, the scalloped shape being produced by suddenly introducing the electron beam into a longitudinal magnetic field at other than the correct equilibrium radius.

Haeff[128, 129] showed that amplification can be obtained if the electrons of two beams, travelling with different velocities, are intermingled. It is necessary[9, 166] that the velocity distribution of the combined beam should possess distinct peaks: the condition for amplification is that the slow wave on the faster beam should have nearly the same phase velocity as the fast wave on the slower beam.

Double-stream tubes may be analysed[146, 242, 269, 278] by letting the beams have average velocities v_{0A}, v_{0B} and plasma frequencies ω_{pA}, ω_{pB}. The propagation coefficient is given[133] by

$$-1 = \frac{\omega_{pA}^2}{(j\omega - v_{0A}\gamma)^2} + \frac{\omega_{pB}^2}{(j\omega - v_{0B}\gamma)^2} \quad . \quad . \quad (34)$$

The solution of eqn. (34) is of the form

$$\gamma = \alpha + j\omega/v_{0m} \quad . \quad . \quad . \quad . \quad . \quad (35)$$

where v_{0m} is the arithmetic mean velocity of the two beams. Let

$$v_{0A} = v_{0m} + \delta \text{ and } v_{0B} = v_{0m} - \delta \quad . \quad . \quad (36)$$

so that, if $\delta \ll v_{0A}$ or v_{0B}, $\omega_{pA} = \omega_{pB} = \omega_p$. Eqn. (35) becomes

$$\frac{\alpha v_{0m}}{\omega_p} = \pm j\left\{\left(\frac{\omega\delta}{\omega_p v_{0m}}\right)^2 + 1 \pm \left[4\left(\frac{\omega\delta}{\omega_p v_{0m}}\right)^2 + 1\right]^{1/2}\right\}^{1/2}$$

which is seen to represent four possible waves. A plo eqn. (37) shows the dependence of the propagation coeffic of the four waves on the average beam velocity, ch density, velocity separation and frequency. If the paran $\omega\delta/(\omega_p v_{0m}) > \sqrt{2}$, all four values of α are imaginary, representing unattenuated waves, two of which travel faster v_{0m}, two slower. For larger beam currents, when $\omega\delta/(\omega_p$ $< \sqrt{2}$, two of the possible values of α are real, one neg and the other positive. When α is negative, the wave g exponentially with distance along the beam. Further detai double-stream amplification have been given[214, 252, 254, 263] experimental observations reported.[3]

(3.2) Travelling-Wave Tubes

Travelling-wave tubes are assumed here to result from interaction of an electron beam with the forward-trave fundamental space-harmonic of a periodic or slow-wave s ture. Such tubes were first described by Kompfner[185–188] Pierce,[266, 271, 272] since when their detailed operation has extensively analysed.[81, 200, 201, 202, 316, 335, 340] A typical exar shown in Fig. 10, employs a helix as the slow-wave struc the phase velocity is about 0·1c, corresponding to a beam po tial of 2·5 kV. Such electron tubes, in which the only mag field present is that required for focusing the beam, are us designated[356] O-type.

Combination of the propagation relations for the elec beam and the slow-wave circuit gives, for the case of s space charge, the basic travelling-wave-tube equation[133, 30]

$$(\gamma^2 + \beta_c^2)(\gamma - j\beta_e)^2 = -2\beta_c(C\beta_e)^3 \quad . \quad .$$

where

$$C = \left(\frac{E_z^2}{2\beta_e^2 P} \times \frac{I_0}{4V_0}\right)^{1/3} \quad . \quad . \quad . \quad . \quad .$$

is the gain parameter.[272] The quantity $E_z^2/(2\beta_e^2 P)$, relating axial electric field to the power transmitted along the cir has the dimensions of an impedance.

Eqn. (38) is of the fourth power in γ and therefore lea four possible propagation coefficients. In terms of theory,[62, 299, 300, 301] the solutions correspond to fast and space-charge waves and forward- and backward-travelling ci

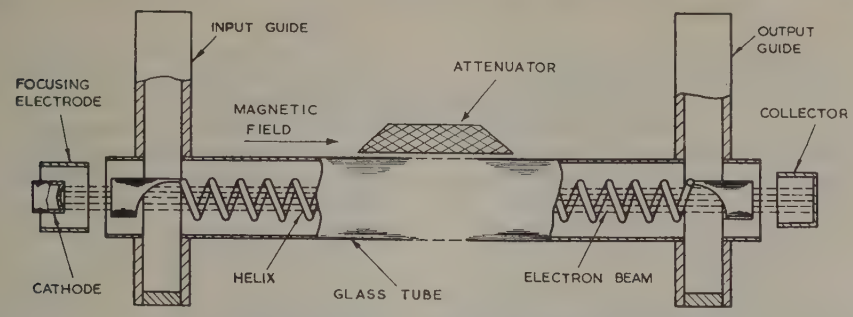


Fig. 10.—Helix-type travelling-wave tube.
Propagation along the helix is from left to right.

waves. In operation of the tube, approximate synchronism obtains between the first three waves, while coupling to the fourth is negligible. On interaction between the forward circuit wave, which has positive power flow, the slow space-charge wave, which has negative power flow, splits into a conjugate pair as in the double-stream tube. Solution of eqn. (38) shows that one wave decreases with distance, the amplitude varying as $\varepsilon^{-0\cdot866\beta_e Cz}$. The other has an amplitude variation of $\varepsilon^{0\cdot866\beta_e Cz}$ corresponding to a gain of $47\cdot3C$ decibels per wavelength. This growing wave forms the output available at the end of the helix. An equal negative power existed in the slow wave and thus the electron beam is finally left with less power than it possessed on entering the helix. The useful output of the tube thus comes from the kinetic energy of the electrons.

Interaction with the circuit wave results in the fast wave, which has positive power flow, varying in phase only, according to the quantity $\varepsilon^{-j(1-C)\beta_e z}$. However, with certain values of beam voltage and current, the circuit wave couples to, and is transferred[119] to, the space-charge wave. Under this Kompfner[189] dip condition, the travelling-wave tube produces no output. The power flow along the tube is constant, the energy being carried by the circuit near the input end and by the electron beam near the collector end.

More detailed calculations of gain have been made[73, 193] and, in particular, the effect of space charge has been examined[104, 109, 247, 336] by the inclusion of Q, a dimensionless space-charge parameter[272] which relates the capacitive impedance between the beam and circuit to the resistive impedance of the circuit. These results enable the gain to be calculated as a function of beam current and geometry. Studies have also been made on the effects of velocity distribution[359] and other factors[46, 195, 199] on gain.

Analysis of the large-signal behaviour of travelling-wave tubes has been made with a view to estimating performance at large power output and efficiency.[86, 245, 362] Partial large-signal theories have been presented,[40, 245] assuming small gain and negligible space charge and attenuation. More refined theories have included effects due to space charge for both small[328] and finite[329] values of C and to other parameters.[165, 235] Results of computations have been plotted[298] graphically to show the r.f. voltage amplitude along the tube length for various values of QC, relative injection velocity and input-signal level. A study of these curves suggests means of improving the maximum saturation efficiency, such as by using a phase-changing device or tapering the helix pitch in order to reduce the wave velocity.

In the travelling-wave amplifier, there is a direct electrical connection from the output to the input terminals. Mismatches can set up a backward circuit wave, resulting in multiple reflections which cause fluctuations in gain with frequency[66] or, in extreme cases, instability and self-oscillation. Such effects can be reduced by introducing attenuation[173] into the circuit: it can be shown that for lumped attenuation[6, 189] only about one-third the inserted value is effective in reducing the forward gain. Experimental investigations have shown[74] how power capacity is limited by attenuator configuration, low resistivities in some attenuators loading[49] the circuit to reduce gain and power output. The effect of attenuation on tubes of high output has been studied by probe techniques.[50]

Non-reciprocal attenuation, such as that based on electronic methods[305] or magnetized ferrites,[67] has the advantage that the forward gain is not impaired. A large effective attenuation can be obtained by the use of two separate helices[227] which are coupled by the velocity modulation and r.f. beam current which persist in the intervening gap. In one example[170, 178] the output helix was operated at a slightly different voltage, a gain of 55–60 dB being obtained at 4 Gc/s.

The extensive number of measurements on travelling-wave tubes include general performance,[348] small-signal characteristics,[306, 307] beam cross-section[308] and the effect of reflections.[171] For investigating some aspects of performance, graphical methods have been employed.[203] Large-signal properties have been investigated[50, 74, 175, 349] including accurate measurements[75] on a tube scaled to a low frequency.

The slow-wave structures employed in travelling-wave tubes have included interdigital lines, sinuous rectangular waveguide[180] and, for frequencies near 10 Gc/s, corrugated coaxial lines.[103] In an application requiring performance over only a 10% frequency band, a tube employing[293] a dispersive helix gave a gain of 20 dB with a beam current of 100 μA. More generally, the helix is used in the non-dispersive region to give bandwidths of one octave or more. This circuit is usually inside the vacuum envelope although external helices have been employed.[225, 226] Auxiliary helices have been proposed[344] for coupling to the input and output ends of the circuit helix and to an external attenuator. If the tube is provided with internal[79, 160] or external feedback, a self-oscillator is obtained. A travelling-wave tube has been proposed[353] as a distributed buncher for a klystron, thus eliminating one tuned circuit: an amplifier of this type could give an output of 10 kW with 35% efficiency, a gain of 18 dB and a bandwidth of 40 Mc/s at 900 Mc/s.

The travelling-wave tube has been mostly employed as an amplifier of low- and medium-level signals and numerous examples have been described.[70, 84, 179, 181, 294, 304] The mid-band gains are in the region 50–60 dB, falling somewhat at the ends of an octave band. Tubes have been developed for $0\cdot7$ Gc/s,[148] 2 Gc/s,[68] 4 Gc/s[296, 303] and $7\cdot5$ Gc/s:[30, 31] an all-metal tube[80] for 3 Gc/s gave a power output of $\frac{1}{2}$ watt for a beam potential of 750 volts and beam power of 15 watts. Lower-

frequency[93] tubes have given 20 dB gain over the band 70–500 Mc/s with efficiencies of 20%. At the other end of the frequency scale, an experimental tube[208] for 50 Gc/s, with a helix of diameter only 0·015 in, gave a gain of 19 dB.

Design information[297] on more powerful tubes has, in conjunction with suitable slow-wave structures, enabled mean powers of 1 kW to be achieved.[360] Recent work[59] has extended the application of travelling-wave tubes to high-power pulse generation. Good thermal dissipation has been achieved[115] by means of an iris-loaded waveguide with slot coupling between successive cavities. At the mid-band frequency of 2·85 Gc/s the gain was 35 dB with an efficiency of 34% and power output of 3 MW. The input was 100 kV, 62 amp, and the $\frac{1}{2}$ in diameter beam was focused with a magnetic field of 1 000 oersteds.

In the transverse[95] travelling-wave tube the electrons are only permitted to travel through the interaction space a distance small compared with its total length. The slow-wave circuit may be a flattened helix, operation being achieved by skewing either the circuit, as in Fig. 11(*a*), or the beam, as in Fig. 11(*b*).

between a phase velocity of v_p for the fundamental, when $n =$ and the value $v_p(d + n\lambda_g)/d$ for a space harmonic of orde The phase velocity of a harmonic required for synchronism v the electron beam can be much greater than that of a fundame wave. Thus, for a given beam velocity and working freque there can be an increased separation of the circuit-elements.

Space-harmonic travelling-wave amplifiers have employe the $n = +1$ harmonic. One such tube contained[228, 229] a rid waveguide loaded with 100 stubs, cut tranversely in the ri spaced 0·534 mm between centres, 0·178 mm wide and $\frac{7}{8}\lambda$ d To obtain good coupling, the electron beam passed along t axial grooves in the ridge. The beam power was 1·2 kV, 4 with a magnetic focusing field of 1 600 oersteds. At a freque of 50 Gc/s the gain was 20 dB and the power output 20 mW.

In backward-wave tubes interaction occurs with the $n =$ or first reverse space harmonic. The phase condition is the

$$v_p d = v_0(\lambda_g - d) \quad . \quad . \quad . \quad . \quad .$$

where v_p is now measured in the reverse direction: the ci

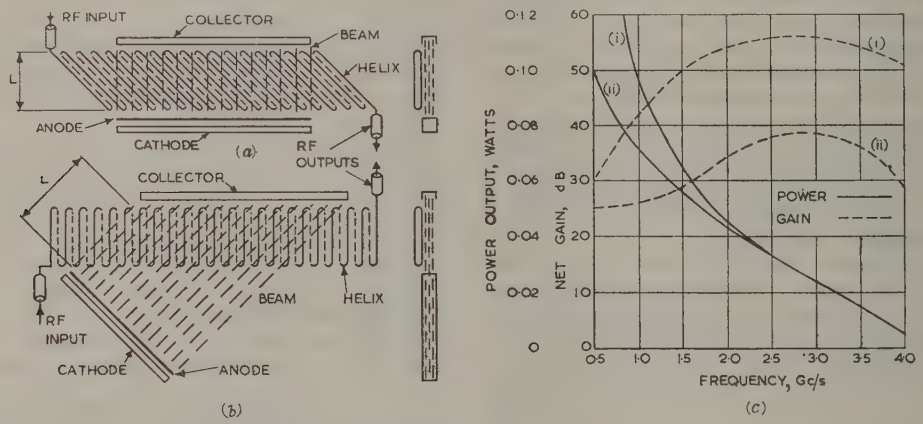


Fig. 11.—Transverse-current travelling-wave tubes.
(*a*) Tube with skew circuit.
(*b*) Tube with skew beam.
(*c*) Theoretical values of gain and power for (i) conventional and (ii) transverse-current tubes.

The properties of the tube are such as to give three forward waves which increase in amplitude exponentially, linearly and as the square of the distance respectively. When the distance is below some critical value, the gain is less than that of the equivalent parallel tube and varies more rapidly with current and frequency. A useful feature of this device is that the power output reaches a saturation value as the input signal is increased, which is dependent upon the performance of the individual beam elements. Experimental transverse-current tubes have shown[94, 101] properties agreeing well with the theoretical values given in Fig. 11(*c*).

(3.3) Backward-Wave Tubes

An electron beam can move in synchronism with, and interact with, one of the higher space harmonics of a periodic structure. In the case of a line loaded with stubs, as shown in Fig. 12, the condition for the electric field at the mouth of each of the stubs to be seen in the same phase by an electron is given by

$$v_p d = v_0(d + n\lambda_g) \quad . \quad . \quad . \quad . \quad . \quad (40)$$

where $n = 0, \pm 1, \pm 2$, etc., and corresponds to the order of the space harmonic. The electron has no way of distinguishing

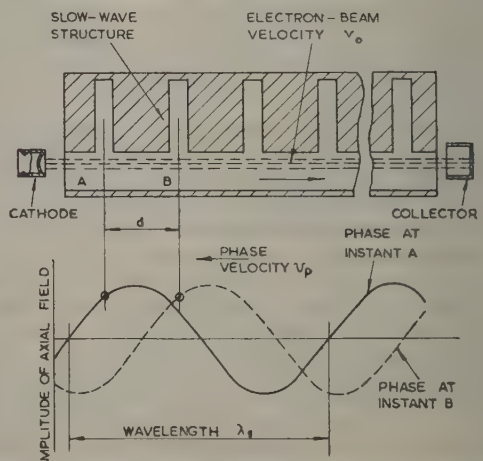


Fig. 12.—Interaction of an electron beam with a space harmon
Drawn for $n = -1$ with v_p in the reverse direction.

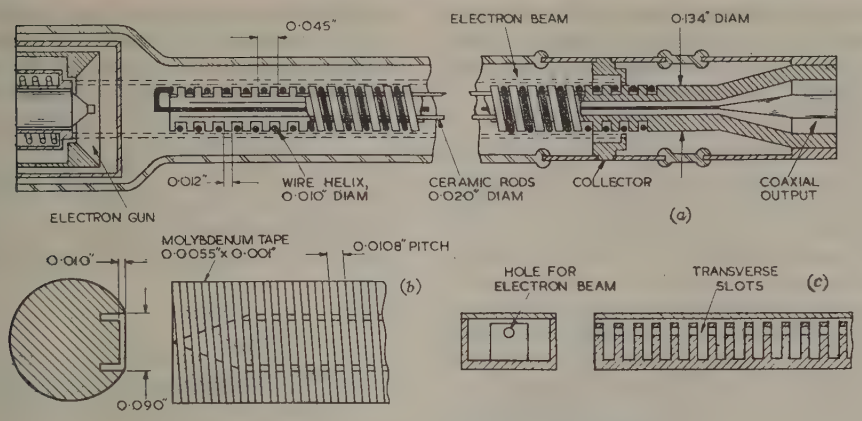


Fig. 13.—Types of backward-wave oscillator.

(*a*) Unbalanced bifilar helical structure with a hollow electron beam.
(*b*) Ridged waveguide with slots in the upper wall.
(*c*) Ridged waveguide with slots in the ridge.

wave and the electron beam travel in opposite directions. This reverse harmonic is well known to be dispersive—the phase velocity varies rapidly with frequency. Since the condition of interaction is approximate synchronism between the circuit and beam velocities, it is evident that the frequency of operation of a backward-wave tube can be varied over a wide range by change of beam voltage. Such tubes[125, 190] are often known as carcinotrons.

The principles[356] of beam-and-circuit interaction follow closely those relating to the travelling-wave tube. Analysis[72, 142] shows that the power/gain ratio can be expressed as $A/(I_0 - I_s)^2$, where A is a constant near unity and I_s is the starting current for oscillation. The gain thus increases rapidly when I_0 exceeds $0{\cdot}9I_s$ and is confined to a narrow band of frequencies. More reliable operation is achieved in the cascade amplifier,[71, 368] in which a first slow-wave structure, such as a helix, modulates the electron stream by interaction with the backward wave and a second helix is, in turn, excited by the modulation on the stream. The advantage over a single-structure tube is increased stability and 'cold' isolation between input and output circuits. The bandwidth, which is normally a fraction of 1%, may be increased somewhat, with a sacrifice in gain, by operating the two structures at slightly different voltages. A third structure, matched at both ends and also synchronized with the backward wave, is found to increase the gain and the gain-bandwidth product over that of the two-structure tube.

Backward-wave tubes have a natural tendency to act as self-oscillators and are usually employed as such. More detailed calculations of the minimum starting current have been made[159, 347] and extended[365] to show that the complex nature of the propagation coefficients for backward-wave interaction can be explained[119] on the coupled-mode theory. The efficiency of these tubes has been investigated[123] on the assumption that the r.f. current amplitude is equal to I_0: it may be shown that, with a helical structure, it is advantageous to use a thin annular beam.

A number of practical oscillators have been developed.[8, 96, 237, 248, 321, 342] The internal arrangement is usually such that the slow-wave structure is joined to the output guide at that end at which the electron beam is injected and, at the other end, to a non-reflecting termination. A reflected wave from a mismatched load represents a reduction in useful power output, but, since the condition for synchronism is not satisfied, it cannot interact with the electron beam and thus does not affect the frequency of operation. In the backward-wave oscillator of Sullivan,[323] shown in Fig. 13(*a*), the slow-wave structure is essentially a coaxial line wound into a helix, part of the outer conductor being cut away to allow interaction with the hollow electron beam. The outer helix is made by machining a groove in the inner cylinder. The inner wire is insulated from this groove by three dielectric strips and is brought back along the axis to the coaxial output. With a magnetic focusing field of 1 000 oersteds the tube operated from a frequency of 13·3 Gc/s, at a beam voltage of 3 kV, to 2·6 Gc/s at 40 volts, the output being in the range 50–5 mW.

Other tubes[232, 373] have oscillated over the range 8–15 Gc/s. An example[158] for 8–18 Gc/s employed a tungsten-tape helix supported by three quartz rods as the slow-wave circuit in order to achieve adequately high impedance, low manufacturing cost, good match to the waveguide output and easy control of microphonic effects. The tape thickness was the minimum consistent with reasonable mechanical strength and the tape width was approximately half the helix pitch for maximum impedance of the reverse space harmonic. The helix diameter was chosen to be as large as possible consistent with avoiding anomalous tuning effects.[317] The helix pitch was then chosen to make the high end of the band correspond to a beam voltage of about 2 kV. A hollow electron beam was employed and the main practical data are given in Table 1.

Table 1

BACKWARD-WAVE OSCILLATOR DATA

Frequency range, Gc/s ..	7·6–12·4	11·6–19·0
Tape width, in	0·019	0·013
Tape thickness, in	0·003	0·002
Helix inside diameter, in ..	0·134	0·092
Helix pitch, turns/in ..	24·0	35·5
Beam outside diameter, in ..	0·122	0·081
Beam thickness, in	0·010	0·004
Beam current, mA	10·0	6·6
Active helix length, in ..	2·15	1·60

Backward-wave oscillators have been constructed[168] with a bifilar helix, including an example[93] for the frequency range 180–780 Mc/s. Tubes developed[249] for the ranges 1–2 Gc/s, 2–4 Gc/s and 6–11 Gc/s employed interdigital lines, measure-

ments of their performance indicating that the basic ideas of backward-wave interaction may be correctly applied to their design and construction. An array of hairpin-like elements has been successfully employed[144] as a slow-wave structure.

These oscillators have been found particularly successful in the millimetre-wavelength region. The main problems in such tubes are the mechanical realization of delay structures of very small size which allow the thermal dissipation of relatively high powers. A backward-wave oscillator giving 1–10 mW in the frequency range 70–50 Gc/s employed[61] a helix of tape, size 0·002 in by 0·005 in, wound with a 0·025 in pitch. Similar tubes have covered the ranges 17–41 Gc/s[244] and 20–40 Gc/s.[122] Karp[162, 163] employed the structure shown in Fig. 13(*b*), which consists of a ridged waveguide loaded with a series of transverse slots cut in the broad wall opposite the ridge. This slotted wall may more easily be prepared by winding a grid of gold-plated molybdenum tape around the waveguide, or alternatively by photo-etching a thin metal foil. The electron beam traverses the central portions of the slots, where the axial electric field is a maximum, while the slots are, with the ridge, gradually tapered away at the output end to present a match to the waveguide. Oscillations were obtained at 57–61 Gc/s by adjustment of beam potential to 900–1 170 volts. Other similar tubes have been made to operate at frequencies up to 200 Gc/s. In the arrangement[343] shown in Fig. 13(*c*) the electron beam travels through holes in a series of vanes: several such tubes have covered the frequency range 15–100 Gc/s. A similar vane structure was employed[190] in a tube which oscillated at 40–51 Gc/s as the beam potential was varied in the range 1·6–4·3 kV.

Part 2. CROSSED-FIELD TUBES, ULTRAMICROWAVE GENERATION AND ELECTRON SOURCES

(4) INTERACTION IN CROSSED FIELDS

(4.1) Amplification

Tubes in which there is a field normal to the electron motion are capable of realizing high efficiency, since it is the potential energy of the electrons that is converted into r.f. energy: the kinetic energy is unchanged and thus interaction can be maintained with a slow-wave circuit over a long period. In the E-type tube[477, 603] an electron beam flows between the walls of a parallel-plane transmission line which is bent in the form of a ring. A transverse electrostatic field applied between the two conductors causes the beam to follow a circular trajectory in which the centrifugal and electrostatic forces are equal and opposite. The inner conducting surface supports a slow wave with which the beam interacts to give amplification.

To obtain satisfactory phase focusing of E-type tubes, it is necessary that the number of electronic wavelengths per revolution should be of the order $\sqrt{2}$. A longer interaction period between the beam and the circuit wave can be achieved by giving the beam an initial axial drift velocity so that it spirals along the structure. This principle is employed in the spiratron[595] in which, however, the r.f. interaction involves only the axial component of the circuit electric field and not the angular component. The helitron[619] can be visualized by considering a beam of electrons travelling in a helical path encircling the circuit, which may be a four-conductor transmission line. The r.f. interaction is between the electrons and the angular and radial components of the circuit electric field, and thus the tube is an E-type; no magnetic field is required. One example oscillated between 1·2–2·4 Gc/s with a tuning potential of 650–1 700 volts. The starting current was 0·4 mA for a structure 4 in long.

Crossed electric and magnetic fields are more usually employ and this so-called M-type interaction enables the equivalent the O-type tubes to be realized. If the beam is enclosed walls having infinite admittance the only fields within the be are those due to space charge. Growing waves are always up[405, 467, 611, 618] since only a monotonic variation of electr velocity is required with crossed fields. Under certain con tions, growing waves may exist[432] if the wall admittance resistive or reactive components.

Travelling-wave interaction[401, 429, 430] occurs in the pla structure[491, 610] of Fig. 14(*a*), the electron beam flowing betw

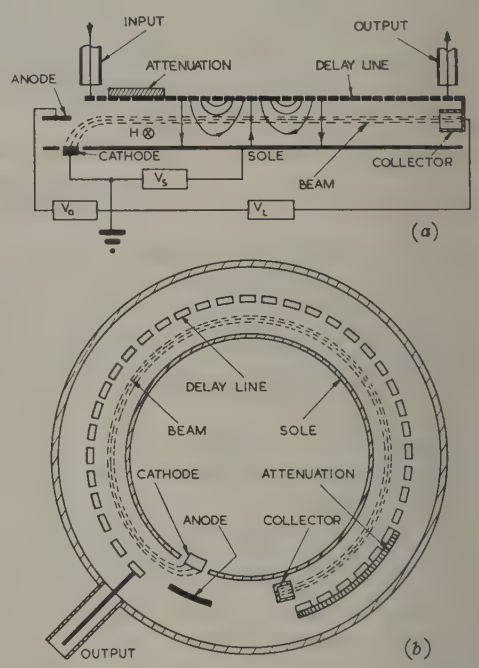


Fig. 14.—Crossed electric- and magnetic-field tubes.
(*a*) Planar version of the T.P.O.M. tube.
(*b*) Circular version of the 'M'-carcinotron.

a negative electrode, termed the sole, and a slow-wave or de structure. There is a constant magnetic field perpendicular the direction of constant electric field and to the motion of electrons. If the delay line and sole are a distance d apart, V the voltage between them and

$$v_0 = V_L/Bd \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad ($$

then the path of the electron beam is parallel to the electro

When a wave is propagated along the delay line under c ditions of synchronism, the electrons interact to give ph focusing.[629] This process provides a mechanism by wh electrons that have absorbed energy have their paths cur towards the sole so that they are removed from the interact space. Electrons which give energy to the circuit wave curved towards the delay line, while being maintained in correct phase of the r.f. field, and form into discrete bunches

Under small-signal conditions and with negligible space cha the increase, dI_1, of the alternating current between plane and $z + dz$ along the beam must be proportional to the electric field and to the length dz, so that

$$dI_1 = KE_z dz \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad ($$

The variation, dP, in the power flowing along the line is that due to the power gained by the beam. The latter is proportional to $E_z I_1 dz$, and since P is proportional to E_z^2 and dP to $E_z dE_z$,

$$dE_z = K' I_1 dz \quad . \quad . \quad . \quad . \quad . \quad . \quad (44)$$

Elimination of I_1 or E_z from eqns. (43) and (44) gives

$$\frac{d^2E_z}{dz^2} = \gamma^2 E_z \text{ and } \frac{d^2I_z}{dz^2} = \gamma^2 I_1 \quad . \quad . \quad . \quad (45)$$

where $\gamma^2 = KK'$. More complete calculation shows that at synchronism γ^2 is real and is given by

$$\gamma^2 = \frac{Z_c \omega d}{Z_e v_p} \coth \omega\delta/v_p \quad . \quad . \quad . \quad . \quad (46)$$

where δ is the distance between the beam and the sole. The value of Z_c is given by

$$Z_c = \frac{E_z^2 v_p^2}{2\omega^2 P} \quad . \quad . \quad . \quad . \quad . \quad . \quad (47)$$

The field distribution along the line, putting $I_1 = E_z = 0$ when $z = 0$, is given by

$$I_1 \propto \sinh \gamma z \text{ and } E_z \propto \cosh \gamma z \quad . \quad . \quad . \quad (48)$$

and γ becomes the gain coefficient of the tube. These results are modified under large-signal conditions.[444]

The efficiency may be estimated by assuming that the electrons have initially a rectilinear flow along an equipotential, V_0. In the presence of the r.f. field, an electron moves towards the anode along trajectories orthogonal to the lines of force. The electron receives from the constant electric field an energy $e(V_L - V_0)$. Since its kinetic energy remains unchanged and equal to eV_0, the energy $e(V_L - V_0)$ is transferred to the r.f. field. Under these conditions

$$\eta_e = 1 - \frac{V_0}{V_L} \quad . \quad . \quad . \quad . \quad . \quad . \quad (49)$$

In fact, not all the electrons reach the anode, and the above efficiency should be multiplied[614] by a factor of 0·8. The efficiency is also reduced by beam interaction[627] with interfering space harmonics and by space-charge wave amplification.[612] Another source of reduced efficiency is cycloidal motion of the electrons, which gives them rotational energy not normally available for conversion into r.f. energy. The electrons in their cycloidal motion reach the equipotential, $4V_0$, and the energy available for transfer is only $e(V_L - 4V_0)$. The electronic efficiency thus becomes

$$\eta_e = 0{\cdot}8\left(1 - A\frac{V_0}{V_L}\right) \quad . \quad . \quad . \quad . \quad (50)$$

where A lies between 1 and 4, depending on the electron trajectories.

The plane structure considered is not, of course, necessary, and a tube in which the delay line is bent to circular shape is easier to manufacture and requires a more compact magnet. In one typical tube,[613] an inter-digital delay line was used in order to achieve low dispersion and high coupling impedance. At frequencies of 1·2–1·5 Gc/s the gain was 10–20 dB, the efficiency was 35–40% and the power output was 200–450 watts. The crossed-field principle can also be applied to high-power pulsed amplifiers, and, for example, a peak power of 1 MW, 20 dB gain, 40% efficiency should be possible at a frequency of 3 Gc/s with a bandwidth of 12%.

(4.2) M-Type Carcinotrons

A backward-wave tube operating with a crossed magnetic field results in the M-type carcinotron.[391, 435, 538] The power flows in a backward direction along the delay line: dP is the negative of the power given by the beam between z and $z + dz$, and thus a negative sign must be inserted in eqns. (44) and (45). The boundary conditions are $I_1 = 0$ at $z = 0$ and $E_z = 0$ at $z = l_c$, the circuit length. The gain conditions become

$$I \propto \sin \gamma z \text{ and } E_z \propto \cos \gamma z \quad . \quad . \quad . \quad . \quad (51)$$

The starting current for oscillations is obtained, since $\cos \gamma l_c = 0$, from the relation

$$\gamma^2 = \frac{Z_c \omega d}{Z_e v_p} \coth \frac{\omega\delta}{v_p} = \frac{\pi^2}{4l_c^2} \quad . \quad . \quad . \quad . \quad (52)$$

Space-charge conditions give rise to the diocotron effect,[612] which reduces the starting current. The electronic efficiency is approximately that given in eqn. (50).

These backward-wave tubes may exist in planar or, as shown by the example of Fig. 14(*b*), circular form. The frequency may be varied by means of the magnetic field or, as is more usual, by the line voltage. The tuning band is limited either by the cut-off frequency of the delay structure, or by practical considerations such as maximum voltage or maximum direct-power input. If the variation of power output is not a consideration, very wide tuning bands may be achieved, the variation of frequency with voltage being smooth and almost linear over the whole range. This characteristic can be used at a very high speed, and, therefore, for frequency modulation.

The M-type carcinotron has proved useful for high-power high-efficiency operation.[613] In one example, the power output was about 400 watts with an efficiency of 40%, the frequency being variable over the range 1·75–2·85 Gc/s with an applied voltage of 2·6–5·8 kV. As with other backward-wave structures, the frequency of oscillation is insensitive to load impedance and the lines of constant output on a Rieke diagram would be nearly circular. In a coaxial design[617] for pulse operation, the magnetic field is produced by a large current passing along the inner conductor. Other M-type wide-band tubes include the carmatron oscillator[616] and the platinotron amplifier.[402]

(5) MAGNETRONS

(5.1) Regimes of Oscillation

One large class of microwave tubes, depending for operation on crossed electric and magnetic fields, is the magnetron. The basic structure of periodic circuit and electron beam is bent around in a circle so that the path of amplification is a closed loop: the tube supplies its own input and thus becomes an oscillator. The cylindrical sole takes the form of a thermionic cathode, so that electrons are emitted along the tangential length of the structure. From Fig. 15(*a*), the equations of motion of an electron are

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = -\frac{erB}{m}\frac{d\theta}{dt} + \frac{e}{m}\frac{dV}{dr} \quad . \quad . \quad (53)$$

$$\frac{1}{r}\frac{d}{dt}\left(r^2\frac{d^2\theta}{dt^2}\right) = r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt} = \frac{eB}{m}\frac{dr}{dt} \quad . \quad . \quad (54)$$

Rearrangement and integration of eqn. (54) gives

$$\frac{d\theta}{dt} = \frac{eB}{2m}\left(1 - \frac{r_c^2}{r^2}\right) \quad . \quad . \quad . \quad . \quad (55)$$

When the electron is at its most distant point from the cathode, the radial velocity is zero and thus

$$\left(\frac{d\theta}{dt}\right)^2 = \frac{eV}{mr^2} \quad . \quad . \quad . \quad . \quad . \quad (56)$$

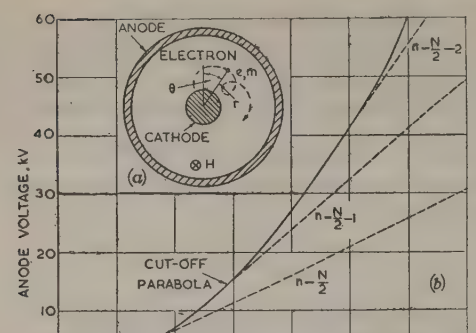


Fig. 15.—Anode voltage and magnetic field in a magnetron.
(*a*) Geometry of magnetron.
(*b*) Cut-off parabola and Hartree lines.

If the electron just grazes the anode, the relation

$$V_a = \frac{er_a^2}{8m}\left(1 - \frac{r_c^2}{r_a^2}\right)^2 B^2 \quad . \quad . \quad . \quad . \quad (57)$$

gives the cut-off parabola of Hull[481] shown in Fig. 15(*b*). Space-charge wave amplification and other effects are thought to prevent the cut off from being abrupt and complete in practice.[475, 604]

The anode current of a diode is given by the $V^{3/2}$ relation of Langmuir[503] up to the onset of saturation, but it is modified in the presence of a magnetic field to become

$$I = \frac{\epsilon_0 r_c^2 e^2}{m^2} A\left(\frac{r_a}{r_c} - 1\right) B^3 \quad . \quad . \quad . \quad (58)$$

where A is a slowly varying function of r_a/r_c. Typical curves are shown in Fig. 16. If the cathode radius is small and space-charge effects are negligible, the electron moves in an approximately circular orbit with angular velocity eB/m so that

$$\lambda B = 2\pi mc/e \quad . \quad . \quad . \quad . \quad . \quad . \quad (59)$$

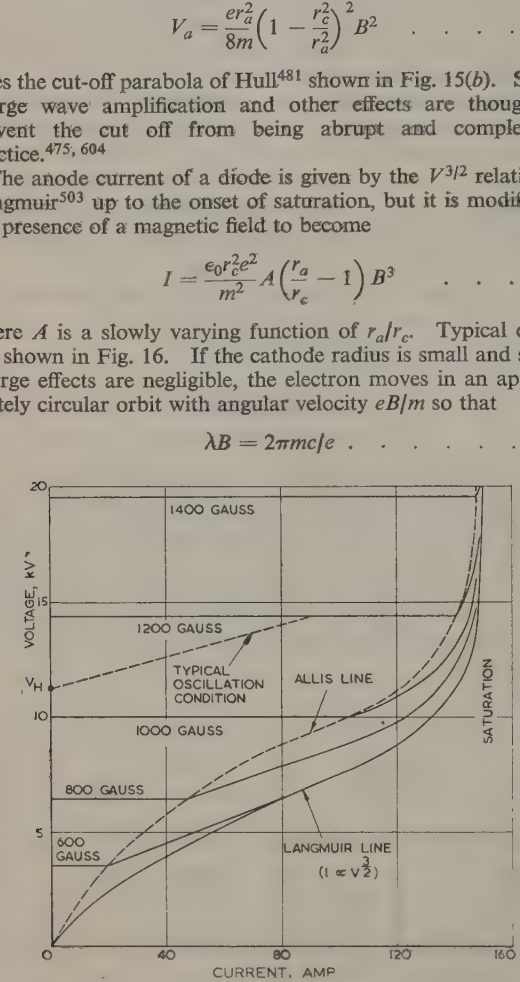


Fig. 16.—Voltage/current characteristics of a magnetron.
The typical oscillation condition for 1 200 gauss cuts the vertical axis at the Hartree voltage.

R.F. oscillations corresponding to eqn. (59) are observed[473, 476] on connecting a resonant circuit between anode and cathode, w negative-resistance low-frequency oscillations are obtained[474] when the anode is divided into one or more pairs of segme between which a resonant circuit is connected.

(5.2) Electrical Circuit Structure

At microwave frequencies the magnetron anode takes the fo of a periodic structure, to a space-harmonic of which the elect beam couples. Since this circuit is a complete circle, only wa of those frequencies which result in an integral number of wa lengths around the closed path can exist. The structure r be interdigital,[482, 484, 507, 509, 510] but more usually takes form of a strip transmission line or waveguide series-loaded v stubs, as shown in Fig. 17. The axial-field variation is z for a transmission line and half-sinusoidal for a waveguide, w the only radial mode of importance is that giving a quar period variation from the mouth of the stub.

The circumferential modes are more complicated and fo stubs and n cycles of field variation around the anode, n r have the values 0, 1 . . . $N/2$. All modes except $N/2$ are dou degenerate, allowing travelling waves to propagate in ei direction. The $N/2$ mode is a singlet and gives a standing w in any case. There are, moreover, higher space harmonics giv additional cyclical variations of electric field. In general, number of repeats is given[415] by

$$m = n + pN \quad . \quad . \quad . \quad . \quad .$$

where p is any integer, positive or negative, or zero. Nega values of m indicate components travelling in a direction oppo to the fundamental.

If, for example, N is 8, the phase differences for the fun mental modes, $p = 0$, $n = 1$, 2, 3 and 4 are $\pi/4$, $\pi/2$, 3 and π. These relations are illustrated in Fig. 18, where anode structure has been opened out.[446] For the case of n = the space harmonic with $p = -1$ is also shown. The ac calculation of the resonant frequencies is difficult, and accu results have only been given[396, 508] for infinitely long anodes

The separation in frequency between, say, the π-mode the unwanted modes may be increased by strapping. straps consist[415, 550] of thin wire or strip connecting alter segments, and, as shown in Fig. 19, may exist in many for The usual symmetrical forms of strapping are single-ring, eche and double-ring. Unsymmetrical strapping[600] may be achie by inserting breaks or by employing the YB-system. For mo other than the π-mode, current flows in the straps, the inducta of which detunes the circuit. The π-mode frequency is chan slightly and may, in fact, be preset within limits by adjustn of the straps. The improvement due to either single- or dou ring strapping of a typical 8-stub anode is shown in Table 2.

Mode separation at frequencies higher than about 10 Gc/

Table 2

EFFECT OF STRAPPING AN EIGHT-STUB ANODE

Mode number	Resonant frequency		
	No strapping	Single-ring	Double-ring
	Gc/s	Gc/s	Gc/s
4	3·00	3·00	3·00
3	3·03	3·16	4·68
2	3·12	4·14	6·38
1	3·30	4·86	8·11

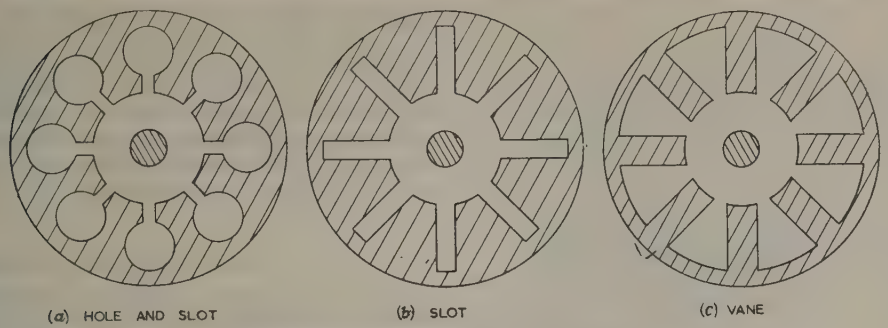


Fig. 17.—Types of magnetron anode structure.

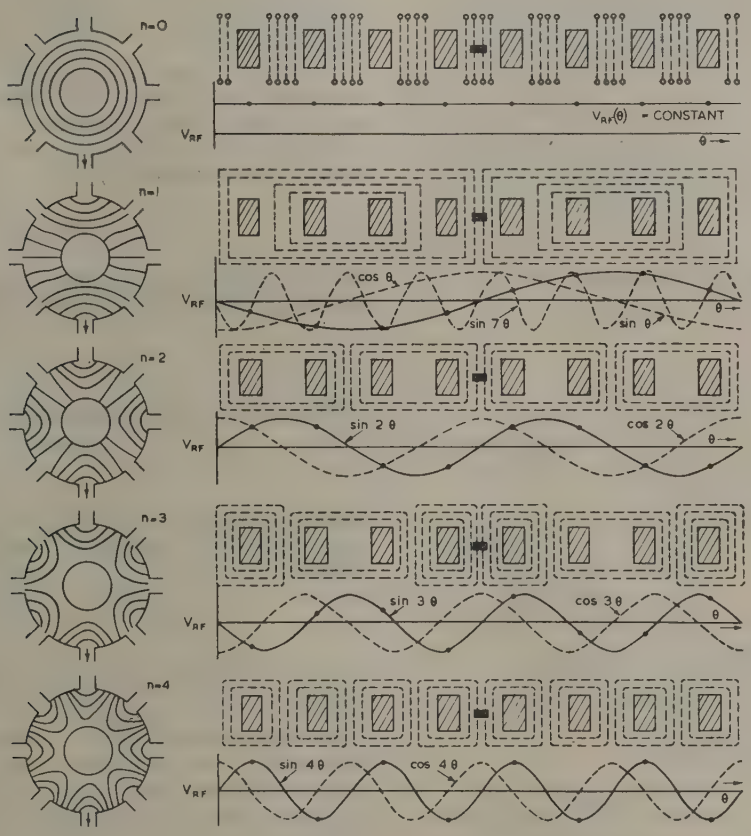


Fig. 18.—Mode patterns in a magnetron.

The full lines represent electric field and the dotted lines represent magnetic field.

nore easily achieved by systematic modification of the slow-wave tructure. In one method,[499, 530] the stubs are alternately long nd short. A typical example of an 18-stub 'rising-sun' anode s shown in Fig. 20(*a*). The resonant wavelengths of the struc- ure are shown in Fig. 20(*b*) and are seen to fall into two groups. As the length ratio increases, maintaining the π-mode wave- ength constant, the wavelengths of the modes $n = 1$ to $n = 4$ ncrease, while those of modes $n = 5$ to $n = 8$ decrease. The oltage antinode recedes into the long stubs and the net radial ield thus caused produces an interfering $n = 0$ or zero mode. There is thus an optimum length ratio of about 2. When the number of stubs becomes large, improved performance is obtained by partly or completely closing the anode structure with end plates.[480]

Another anode structure makes use[445] of the TE_{011} mode in a coaxial cavity, concentric with an inner resonator assembly. The rear walls of alternate resonators are slotted through to couple the system. The anode can be tuned in frequency by plungers in the coupled guide, the construction being suitable for frequencies of 10 Gc/s. A magnetron structure has been

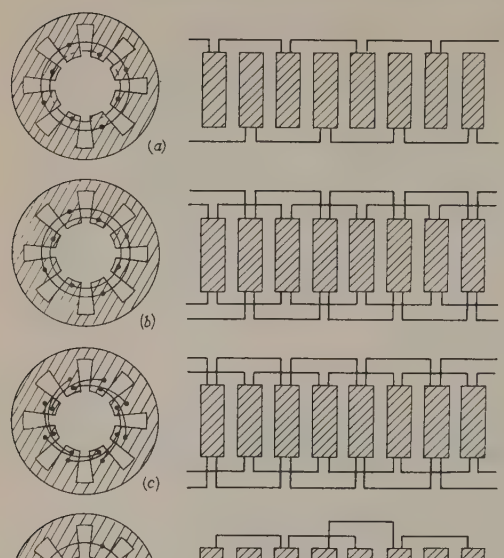
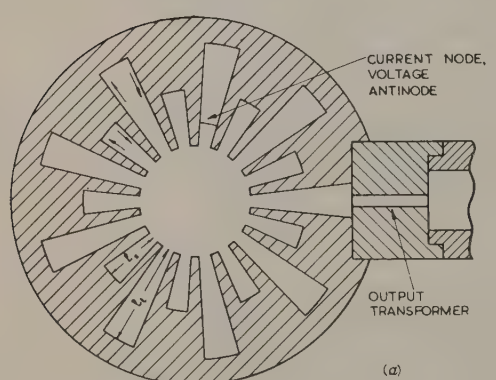


Fig. 19.—Strapping systems.
(*a*) Single ring. (*c*) Echelon.
(*b*) Double ring. (*d*) YB-type.



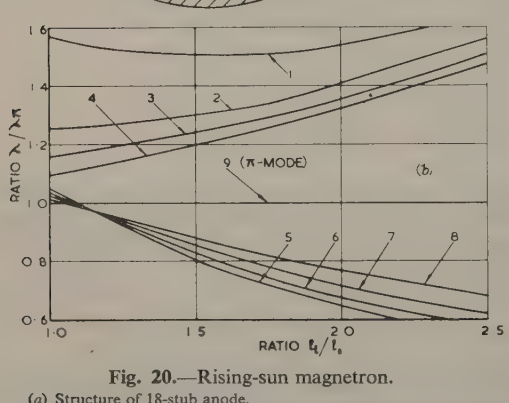


Fig. 20.—Rising-sun magnetron.
(*a*) Structure of 18-stub anode.
(*b*) Mode plot, λ_π being maintained constant by choice of l_1 and l_2.

proposed in which the cathode and anode are concentric toroi the constant electric field is radial, and the magnetic field loo around the cathode azimuthally. The electrons travel along t cathode, at right angles to both electric and magnetic fiel The circuit wave which is excited by them follows the sa direction, and the anode is thus provided with transverse resona stubs. In all these anode structures, the coupling[597] to t external waveguide or transmission line may be by probe, lo or slot excitation.

(5.3) Electron and Circuit Interaction

The interaction space of a magnetron contains a steady elect field between cathode and anode, a uniform axial magnetic fie electrons emitted from the cathode, and an r.f. field betwe cathode and anode, consisting of a wave travelling slowly arou the anode structure and increasing rapidly in intensity as t latter is approached. The solution of this interaction proble is difficult, but some progress has been made with analyti methods.[415, 589]

The effect of a small r.f. electric field is to change the directi rather than the energy of the electrons. For oscillations start, the electric field must be large enough to give the electro a velocity equal to that of the r.f. wave. If the radial accelerati in eqn. (53) is made zero, then

$$\frac{dV}{dr} = v_0\left(B - \frac{2mv_0}{r_a + r_c}\right) \quad . \quad . \quad . \quad . \quad (6$$

assuming the electron is travelling at velocity v_0 midway betwe cathode and anode. As a simplifying approximation let

$$\frac{dV}{dr} = \frac{V_a}{(r_a - r_c)} \quad . \quad . \quad . \quad . \quad . \quad (6$$

The condition of synchronism with the circuit wave gives

$$v_0 = \frac{\pi f}{n}(r_a + r_c) \quad . \quad . \quad . \quad . \quad . \quad (6$$

Substitution of eqns. (62) and (63) in (61) extends the conditi originally given by Posthumous[558] to give the well-known Hart voltage:

$$V_H = \frac{\pi f(r_a^2 - r_c^2)}{n} B\left(1 - \frac{1}{n}\right) \quad . \quad . \quad . \quad (6$$

A typical oscillation condition and value of V_H are shown Fig. 16. A nomogram for this threshold voltage has be given.[471]

Eqn. (64) represents the cut-off condition for the oscillati magnetron. Lines are plotted in Fig. 15 for various values n, the Hartree lines being tangents to the cut-off parabola. the anode potential is increased from low voltages, where curr cut-off obtains, oscillations occur at the $N/2$ line, restart at $(N/2) - 1$ line, and so on until the cut-off parabola is reach when the region of steady current prevents further oscillati These relations permit scaling[594] of a magnetron from one f quency to another.

The exact mechanism of build-up of oscillations obscure.[404, 454] Experimental evidence[475] suggests that dense space-charge cloud surrounding the cathode becon unstable and causes pre-oscillation before the threshold c dition of eqn. (64) is reached. The stability of this cloud been the subject of much study.[453, 473, 478, 540] Inclusion of effect of electron emission velocities predicts[598] a diffusion the cloud edges which is supported[523] by experiment. analysis[532, 591, 599] of the pre-oscillating state, the adoption a statistical approach shows[452, 587] that peaks occur in the nc

pectrum. These have been observed in practice,[461] thus ndicating a smooth transition between the pre- and full-scillating states. Injection of a small signal enables[428] the hase of the oscillations to be controlled.

The electron orbits under full oscillating conditions are ifficult to analyse, but in some instances numerical calculation as been made. Thus, Fig. 21 shows the orbits of the CV76

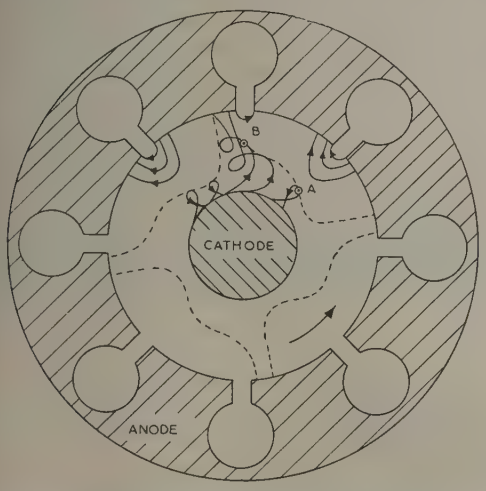


Fig. 21.—Electron bunching in a magnetron.
The r.f. electric field at the anode causes phase focusing of the electrons.

plots[561] four parameters: two associated with the input circuit and two with the output. Input voltage and current are plotted, respectively, as ordinates and abscissae, and on this plane, lines of constant magnetic field and constant power output and/or efficiency are drawn. These curves are plotted for the mode of primary interest and for a fixed external load. If the overall efficiency remained constant, the power-output contours would be hyperbolae that approach the axes asymtotically. Practical charts bear some resemblance to the static characteristics of Fig. 16, as may be expected. An example[415] for an 18-stub rising-sun magnetron operating near 9·5 Gc/s is shown in Fig. 22. The variation of overall efficiency with current, for a constant field, is related to the extent to which the rotating space charge departs from synchronism with the r.f. field on the anode structure. In this particular tube, there is a zero-mode dip near 3 500 gauss which is, however, below the operating field.

The magnetron is well known as a generator of large pulsed power and the development of the original cavity type for 3 Gc/s has been recorded.[395, 525, 567, 568, 569] Tubes were later constructed[415, 469, 624, 625] for frequencies in the range 1–24 Gc/s with power outputs up to 1 MW peak. Further developments have been in the direction of longer life, more arduous duty cycles and improved frequency stability. Recent tubes[495, 544] include a magnetron producing,[545] at 1·31 Gc/s, a peak power of 10 MW with 10 microsec pulse duration and 17 kW mean power. The input was 66 kV, 400 amp, the magnetic field was 1 440 oersteds and the ring-strapped anode had 10 hole-and-slot resonators.

The frequency characteristics of magnetrons have been the subject of much study.[620, 621] A limited range of frequency

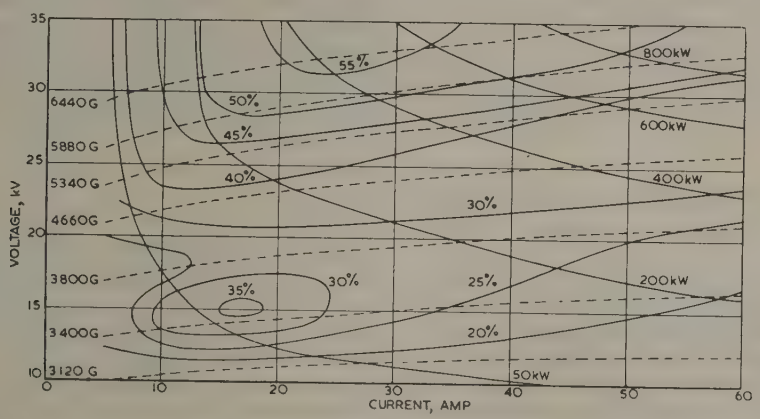


Fig. 22.—Performance chart of a magnetron.

Anode, 18-stub rising sun; frequency, 9·375 Gc/s; pulse length, 0·5 microsec; p.r.f., 500 c/s.
The efficiency dip near 3 500 gauss should be noted.

agnetron operating at a frequency of 3 Gc/s. An electron at oint A gains energy from the r.f. field and returns to the cathode, hile an electron at point B gives up energy and eventually ikes the anode. This bunching of electrons results in the ace-charge cloud taking on[566, 582] the spoke-line appearance own.

(5.4) Practical Data

The performance of a magnetron as a function of load impe-nce may be given in the form of a Rieke diagram. The output upling is normally adjusted so that a matched-load impedance ves the optimum conditions for output and efficiency. The actical aspects may also be represented on a chart which

tuning can be achieved as a result of modification[424, 541] of the slow-wave structure, for example by the simultaneous insertion[415] of conducting pins into the elementary resonators. Greater ranges are achievable with voltage tuning in conjunction with the use of a higher-order space harmonic:[400, 552] such operation is facilitated by the adoption of interdigital structures.[397, 586]

Rising-sun magnetrons are preferred[562, 563, 564] for generating high power at millimetre wavelengths.[571] An example[601] for 35 Gc/s gave 20 kW with a pulling figure of 32 Mc/s: the input was 13 kV, 10 amp and the magnetic field of 10 500 oersteds was below the zero-mode dip. Another tube[630] gave 100 kW peak, 40 watts mean power with a life of 200 hours. Using 22-vane structures,[392] a power of 100 kW at 48 Gc/s has been

obtained with an input of 20 kV, 25 amp: another tube, for 90 Gc/s, gave 10 kW with an input of 10 kV, 10 amp, a magnetic field of 26 000 oersteds and a life of 100 hours.

A different approach[573] to millimetre-wavelength operation has shown that magnetrons can oscillate with reduced efficiency at magnetic fields of about one-quarter of the conventional values. Typical minimum-voltage-regime tubes have covered[437] the range 10–60 Gc/s. Still lower voltages can be employed with spatial-harmonic operation.[572] A typical anode can be thought of as having 12 uniformly spaced resonators, eight of which were omitted in two groups of four. Oscillations were obtained

amplitudes. A well-bunched beam results, in practice, in increase of radiated power by a factor of 10^7 or more and thus always employed.

For a delta-function current pulse, the harmonic-curr amplitude, I_s, is given by $2I_0$. This enables high harmonics the input frequency to be generated with the aid of suita circuits. For example, energy has been detected,[533] correspo ing to the 12th to 65th harmonics, when an electron beam 2 MeV energy and bunched at 2·83 Gc/s was passed axi along a 0·42 in × 0·17 in waveguide. Such harmonic quencies have also been achieved by the arrangement of Fig. 2

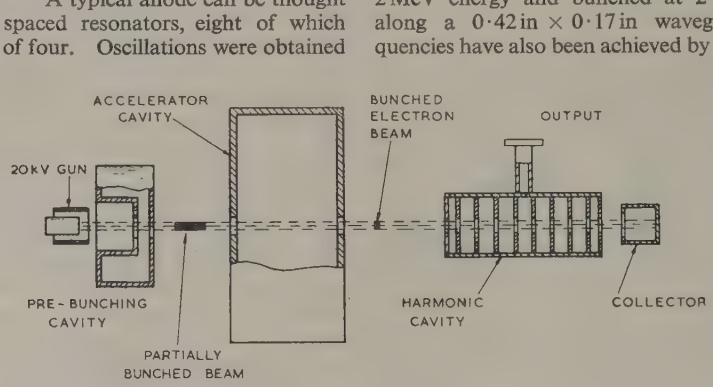


Fig. 23.—The harmodotron—a beam harmonic generator.
Frequency 2·775 Gc/s; beam energy 1–1·5 MeV.
The smaller cavity selects a high harmonic.

by interaction between the electron stream and a component field having the same periodicity as that produced by the full number of resonators. At a frequency of 9 Gc/s, a c.w. power of 75 mW was generated with an input of 600 volts, 9 mA and a magnetic field of 2000 oersteds. A low-field tube for 24 Gc/s possessed[447] a rising-sun structure mechanically tuned with capacitive pins: at a fixed operating point, an output power exceeding 10 watts at an efficiency of 5% was obtained over a frequency band of 7%.

(6) GENERATION OF ULTRAMICROWAVE FREQUENCIES

(6.1) Radiation from High-Energy Beams

Power at frequencies above 0·1 Gc/s may be obtained[385, 386] by direct radiation from relativistic electron beams. If the beam is uniform, the contributions of individual electrons are random in phase and the energy radiated is equal to the sum of the energies of each electron. If the beam is suitably space-bunched, the radiation is coherent and it is the amplitude of the radiation field which is equal to the sum of the individual

the harmodotron.[414, 588] Electrons from a 20 kV gun are passed through a pre-bunching cavity where they are give desired velocity distribution; they are then allowed to drift finally enter the accelerator cavity to emerge as a high-curr density tightly-bunched beam of small cross-section. The be then passes through the harmonic-frequency cavity wh may be iris-loaded, as shown, or a plain-tunable type. theoretical relationship for the generated power is typically the order of

$$P_e^2 = 2{\cdot}7(I_s^4/f) \times 10^{18} \quad . \quad . \quad . \quad .$$

With a 1 MeV, 5 mA beam, bunched at 2·775 Gc/s, a powe 1 watt at 36 Gc/s was obtained. The highest frequency obser was 95 Gc/s, corresponding to the 34th harmonic.

It has been shown[534, 535, 576] that three possible mechani of radiation from beams relate respectively to linear, sinuso and circular motion of the particles. Consider first the radia from an electron of velocity v, shown in Fig. 24(*a*), in mo along AA′, which is received at point P. As long as the velo component $v \cos\theta$ in the direction AP is near the wave veloc

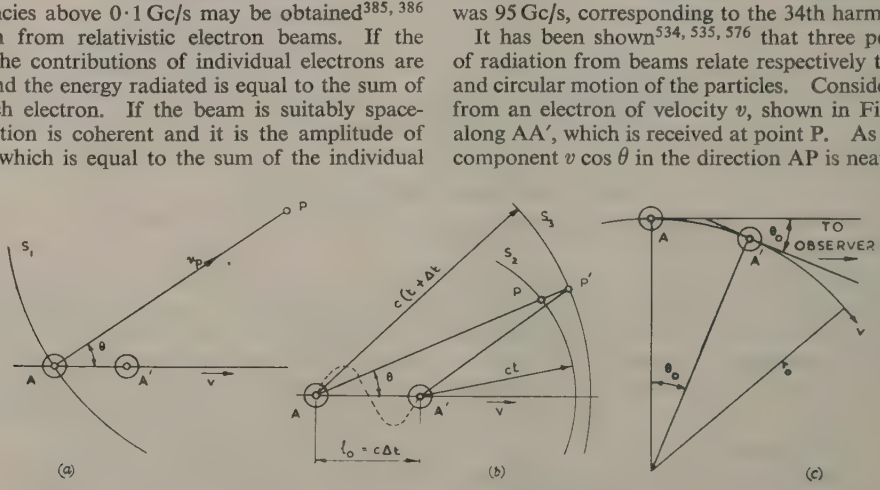


Fig. 24.—Radiation from electron beams.
(*a*) Acceleration in a straight line.
(*b*) Transverse periodic motion.
(*c*) Rotation in a circle.

„ the radiation emitted later catches up with that emitted arlier. The time taken by a spherical wave of potential, S_1, entred at P, to sweep over the charge is a measure of the crease of potential that is observed when this wave finally ollapses on P. This time is proportional to

$$\frac{v_p}{(v_p - v\cos\theta)} \simeq \frac{1}{1-(v/v_p)+\frac{1}{2}(v/v_p)\theta^2} = \frac{2}{(\theta_0+\theta)^2} \quad . \quad . \quad . \quad . \quad . \quad (66)$$

here $\theta_0 = (1 - v^2/v_p^2)^{1/2} = 1/\kappa$. The radiation is contained ithin a narrow cone centred on the direction of movement of ie charge, with angular opening $\theta = \theta_0$.

Such radiation was discovered at optical frequencies by erenkov[407] and its origin was explained by Frank and Tamm.[448] n electron passing, for example, through a dielectric medium, auses polarization charges to be momentarily produced and ie medium reradiates the transferred energy. A wavefront will orm, as in Fig. 25(*a*), in a direction θ such that radiation travels

vector normal to the surface represents radiated energy and is given by

$$S = 4\pi F^2 \rho_0^2 v^2 \varepsilon^{-2qd} \eta_R / c \quad . \quad . \quad . \quad . \quad (70)$$

where

$$q^2 = \omega^2(1/v^2 - 1/c^2) \quad . \quad . \quad . \quad . \quad (71)$$

d is the distance of the beam from the surface

and

$$\eta_R = n^2\left(1 - \frac{v^2}{c^2}\right)\left(\frac{n^2v^2}{c^2} - 1\right)^{1/2} \Big/ 2(n^2-1)\frac{v}{c} \quad . \quad (72)$$

Further, S is given by

$$S = \eta_R Z_w (FA_0)^2 \varepsilon^{-2qd} \quad . \quad . \quad . \quad . \quad (73)$$

The quantity $\eta_R Z_w$ can be interpreted as an intrinsic resistance for Čerenkov radiation in this geometry. For example, if $n = 10$, $v/c = 0{\cdot}2$, $d = 3 \times 10^{-5}$m, $\omega = 6\pi \times 10^{10}$c/s, $A_0 = 10^{-1}$A/m, then, assuming $F = 2$ for perfect bunching,

$$S = 11 \times 10^{-2} \text{ watt/m}^2 \quad . \quad . \quad . \quad . \quad (74)$$

which may be compared with the energy in the 10 kV beam of 1 kW/m.

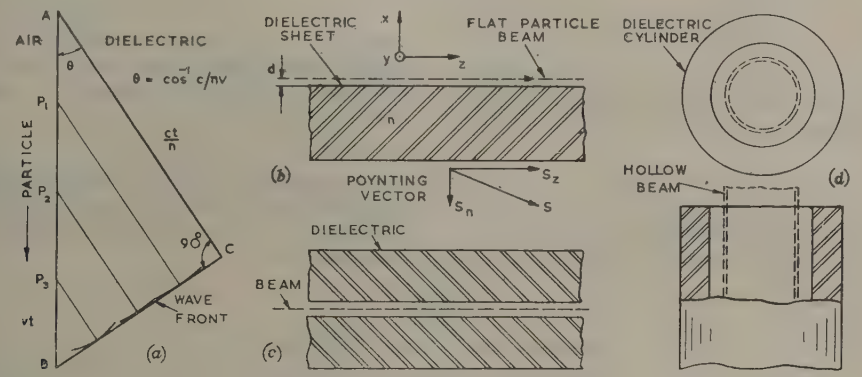


Fig. 25.—Principle of Čerenkov radiation.
(*a*) Radiated wave front.
(*b*) Single flat surface.
(*c*) Cylindrical tunnel.

om A to C. The particle velocity may be greater than the ave velocity, c/n, and thus

$$c/n = v\cos\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (67)$$

efines the conical shell of radiation. The thickness of this shell ecreases as the path of the particle in the medium increases. ince the particle does not lose velocity, the radiation is inde- endent of its mass. The power radiated is given by

$$P = \frac{e^2}{c^2}\int\left(1 - \frac{c^2}{n^2v^2}\right)\omega d\omega \quad . \quad . \quad . \quad (68)$$

here the integral is extended over all frequencies for which $v > c$.

The generation of power at microwave frequencies by the erenkov effect has been considered.[380, 433, 456, 457, 547] The rrangement of Danos[426] requires the particles to move close o a dielectric surface. In the case of a flat beam, shown in ig. 25(*b*), the charge density was assumed to be given by

$$\rho = \rho_0\delta(x)[1 + F\cos(\beta z - \omega t)] \quad . \quad . \quad . \quad (69)$$

here F is the Fourier coefficient of the charge density corre- ponding to the frequency ω. The component of the Poynting

Suitable geometries for Čerenkov generators also include a flat beam sandwiched between two dielectric sheets, as shown in Fig. 25(*c*), a hollow beam in a cylindrical hole[393] in a dielectric, as in Fig. 25(*d*), and other arrangements.[387, 503, 539] The Čerenkov effect resembles in some respects the excitation[516] of surface waves, and corrugated-metal surfaces can function[534] as well as a dielectric.

Practical Čerenkov generators for microwave frequencies[427, 494, 501] have included the use of composite media[575] and anisotropic ferrites.[546] The arrangement of Fig. 26 operated[425] at 24 Gc/s. The radiation was excited by a flat electron beam passing closely to the dielectric material, which was 1·9 cm long and made of polycrystalline TiO_2 with $n = 10{\cdot}2$. The beam potential was 10 kV, the current 0·2 mA and the electrons were previously bunched by passage through a resonant cavity. The radiation contained the bunching frequency and its harmonics: the estimated power in the fundamental was 10^{-7} watt.

The second mechanism of direct radiation from beams depends upon perturbation of the electrons, which can be longitudinal[496] or, as shown in Fig. 24(*b*), transverse. Let S_2, S_3 be two spheres of radii ct and $c(t + \Delta t)$. If an electron takes time Δt to travel the spatial period l_0 from A to A′, the radiation emitted

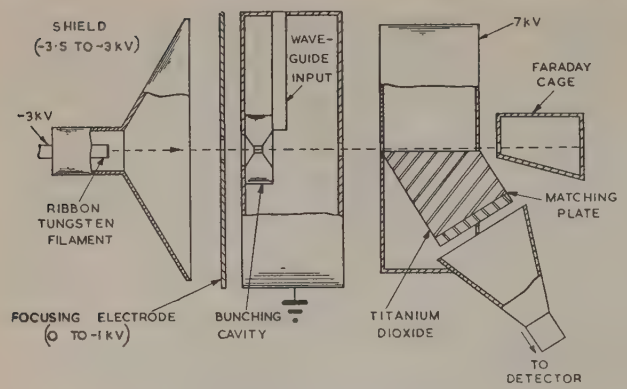


Fig. 26.—Cerenkov microwave generator.
Frequency, 24 Gc/s. Refractive index of dielectric, 10·2.
Output, 10^{-7} watt.

during this interval is piled up in the shell between S_2 and S_3. The thickness PP′ is

$$v_p\Delta t\left(1-\frac{v}{v_p}\cos\theta\right)=\tfrac{1}{2}v_p\Delta t(\theta_0^2+\theta^2) \quad . \quad . \quad (75)$$

If the electron has carried out one period of sinusoidal motion between A and A′, the radiation will contain Fourier components with period $\Delta t/\kappa^2$ or frequency

$$f_{max}=\frac{c\kappa^2}{v_p\Delta t}=\frac{c\kappa^2}{l_0} \quad . \quad . \quad . \quad . \quad (76)$$

This maximum frequency is propagated along the centre of a cone and the radiation tapers off towards lower frequencies for larger angles. It will be seen that the observed frequencies are greater than the frequency corresponding to the spatial period, l_0, because of the relativistic Doppler effect.[502]

The electrons can be given a transverse motion by passage through an undulator—a device consisting of a succession of electric or magnetic fields with alternating polarity. The operating conditions for a magnetic undulator have been studied[416, 524] and a practical model has been constructed.[536] A bunched 5 MeV electron beam was provided by a linear accelerator, and, since the radiated intensity increases with the square of the magnetic field strength, powerful magnets with an array of steel teeth attached to the poles were employed. There were 25 pole pairs with a spacing of 4 cm: the tolerance on the spacing was 0·02 mm and that on the magnet flux 1%. Since the particles travelled at almost the velocity of light the radiated frequencies were given by the intersection of the $v=c$ line in the Brillouin diagram of the periodic structure. In the experiments, small amounts of power, corresponding to 1 watt peak, were observed in the range 0·3–3 Tc/s.

The radiation from a charged particle rotating on a circle of radius r_0 may be examined with the aid of Fig. 24(*c*). While the particle has moved from A to A′, the cone of radiation of angular opening $\theta_0=1/\kappa$ sweeps past the observer and is out of view in a time $r_0\theta_0/v_p$. One such pulse occurs during e revolution and its thickness is $\frac{1}{2}v_p\Delta t\theta_0^2=\frac{1}{2}r_0/\kappa^3$. Hence, h harmonics of frequencies ranging from low values up to

$$f_{max}\simeq c\kappa^3/r_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad ($$

will be observed.

(6.2) Higher-Frequency Limit of Coherent Generation

Methods of generating power at frequencies higher th 0·3 Tc/s have employed the so-called mass radiators[421, 459] wh include black-body sources and spark-type oscillators.[559, Such radiation is incoherent: it resembles noise and cannot p duce independent interference phenomena. In fact, the out of all electron tubes contains some incoherent energy and ther evidence that, according to the type of tube, a higher-freque limit of coherent oscillation exists.

The particular case of interaction between a bunched electr beam and the electromagnetic field in a resonant structure may analysed[434] by including the effect of thermal noise in the cav but neglecting beam noise. The average r.f. beam curr density must be such that

$$J_1^2\geqslant 2\omega^2\epsilon_0 W_{av}/Q_l^2 V_i \quad . \quad . \quad . \quad . \quad ($$

where V_i is the interaction volume of the cavity. W_{av}, average signal energy stored in the resonant field, has a minimu value governed by the requirement that it must exceed the no energy. Thus

$$W_{av}\geqslant h\omega/[2\pi(\varepsilon^{h\omega/2\pi kT}-1)] \quad . \quad . \quad . \quad ($$

For a copper cavity with $T=300°$ K, the values of J_1 for typi microwave tubes at various frequencies are given in Table Thus, for an r.f. beam intensity of 1 A/cm², the frequency lin is about 10 Tc/s.

Table 3

MINIMUM BEAM CURRENT DENSITIES

Frequency, Tc/s	0·01	0·1	1	10	100	1 000
Current density, A/cm² ..	2×10^{10}	8×10^{-7}	10^{-3}	1·5	2×10^{3}	2×10^{6}

A further restriction on coherent oscillation is that the quantu of energy, $h\omega/2\pi$, which is 4×10^{-4} eV at 0·1 Tc/s, becom comparable with the r.f. kinetic and thermal energies of electro in the beam. Since energy exchange occurs only in discre steps of one quantum, such effects modify both the amplitud and phase of the interaction between electrons and the electr magnetic field.

In the particular case of interaction of an electron beam wit a resonant cavity, these quantum-mechanical effects cause th expectation value[578, 579] of the electromagnetic field to diffe from the classicial value by a small amount which, in practic means a small attenuation in signal. Of greater importance the fact that there is an uncertainty or dispersion in this expecta tion value which represents noise in the output signal. Th physical origin of this dispersion lies in three sources. Th first is the quantization of the field[580] and is independent of th electron beam: it represents the uncertainty in the undisturbe electromagnetic field in the cavity resulting from its quantum mechanical properties. The second source of dispersion is i the wave properties of the electron: the fact that it is not a poin causes its interaction to produce an uncertainty in the cavit field. The third source, which is usually of negligible effect, is th classical interaction between the quantum-mechanical uncertaint

the field and the electron, which in turn produced, again in classical manner, an uncertainty in the field.

According to the quantum theory, it is only possible to determine the probability that a given number of quanta have taken part in the interaction with the field. Such effects were first considered[590] for the case of a classical electron and quantized field, and graphical data were given for various examples. This analysis was extended[609] to the case of a quantized electron and classical field, in which, assuming a gap voltage, V_1, the probability of an electron gaining or losing n quanta is $J_n^2(q)$ where

$$q = \frac{2\pi e V_1}{h\omega} \frac{\sin \frac{1}{2}\theta_t}{\frac{1}{2}\theta_t} \quad . \quad . \quad . \quad . \quad . \quad (80)$$

θ_t being the transit angle. In the limit of large number of quanta transferred, this distribution tends to the classical value.

In practice, the increment in field due to a single electron is much smaller than the initial uncertainty in the field. If many electrons are employed it is possible to define a 'signal/noise ratio' given by

$$F = \left(137\pi^2 \frac{V_c}{\lambda^3} \frac{c^2}{v_0^2} \frac{h\omega/2\pi}{\epsilon_0 E_z^2/2}\right)^{-1/2} \quad . \quad . \quad . \quad (81)$$

where V_c is the volume of the cavity. It is seen that this ratio decreases with increase in frequency and increases with electric field. For example, if $V_c \propto \lambda^3$, if the loaded Q-factor is $10^4\lambda^{1/2}$, and if 10^{-4} watt of power is fed into the cavity, then F is unity when the frequency is about 3 Tc/s. The minimum energy in eqn. (78) is then given by one quantum.

Similar calculations[458, 583] confirm these predictions while measurements on the interaction of a beam with a resonant cavity have been reported. The transit time was an integral number of r.f. periods, the resonator being excited at 24 Gc/s by a magnetron. The beam voltage was 10 kV and the beam angle 0·004 rad. The apparent noise temperature of the beam increased after transit and its dependence on the r.f. field indicated quantum effects. In similar experiments,[439] using a transverse r.f. field, the fluctuations in the cavity gave a deflection spread of the electron beam.

(7) ELECTRON SOURCES

(7.1) Thermionic Cathodes

Electron sources in microwave tubes are invariably thermionic cathodes with peak emission densities up to 100 A/cm^2. It is difficult to combine such primary emission with useful life, but back bombardment due to returning electrons assists by causing the emission[487, 488] of a large number of secondary electrons.

Pure metal cathodes are rarely employed, since their emission densities are too low, although water-cooled metal cathodes have been used[602] as pure secondary emitters. The plain-oxide cathode, shown in Fig. 27(*a*), employs a mixture of barium and strontium carbonates supported on a nickel sleeve, and is satisfactory with pulsed emission densities up to 10 A/cm^2. The thorium-oxide cathode of Fig. 27(*b*) is usually employed in high-power tubes. Trouble due to arcing at high voltages may be reduced by packing the active material in the interstices of a woven-nickel mesh, as shown in Fig. 27(*c*). A variation of this cathode is that of Fig. 27(*d*), in which the oxides are held in a sintered matrix of coarse nickel powder. Such cathodes have been employed with success in the 3J50 magnetron operating at a frequency of 10 Gc/s, and the 3J21 for 24 Gc/s.

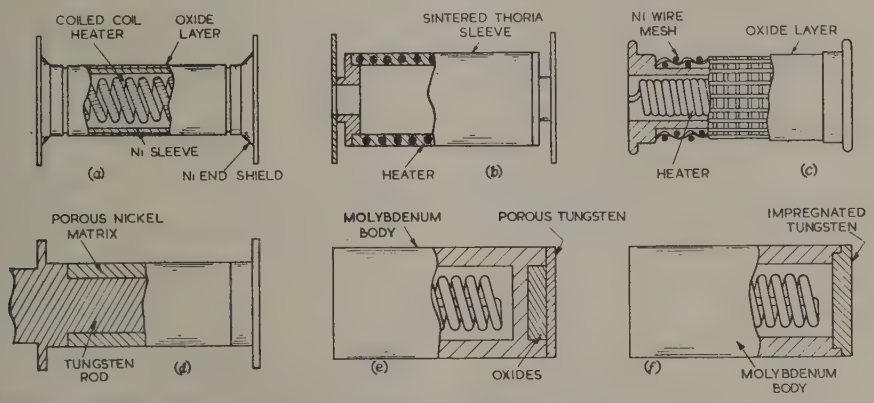


Fig. 27.—Cathodes for microwave tubes.
(*a*) Plain oxide, internal heater.
(*b*) Thorium oxide, internal heater.
(*c*) Nickel mesh, internal heater.
(*d*) Porous nickel, soldering-iron heater.
(*e*) L-type dispenser, internal heater.
(*f*) Impregnated-tungsten dispenser, internal heater.

Recent developments have been associated with the so-called dispenser cathodes. One version, known as the L-cathode,[511] is shown in planar form in Fig. 27(*e*), from which it will be seen that the oxides are contained behind a porous tungsten emitting surface. In the impregnated cathode,[512] shown in Fig. 27(*f*), the alkaline-earth material, in the form of a mixture of normal and basic barium aluminates, is dispersed within the pores of the tungsten body, thus making for a simple construction. Cathodes for microwave tubes may employ a sintered mixture of thoria and molybdenum powders[545] and other materials.[422]

(7.2) Focused Beams

In grid-controlled tubes and magnetrons, the cathode, in conjunction with the positive electric field of the anode, is the source of the electron beam. More usually, the beam is emitted with a circular cross-section by an electron gun. Such guns exist in a number of forms[592] and one example,[554] consisting of a concave cathode and a suitable beam-forming electrode, is shown in Fig. 28.

The subsequent shape of the electron beam is determined by space-charge forces which are usually described by a parameter known as the perveance of the beam or gun. In some tubes, such as low-power reflex klystrons, the beam is short and the effect of divergence may be determined from the current flowing through a cylindrical tunnel. This current is a maximum when the minimum beam diameter is midway along the cylinder. If

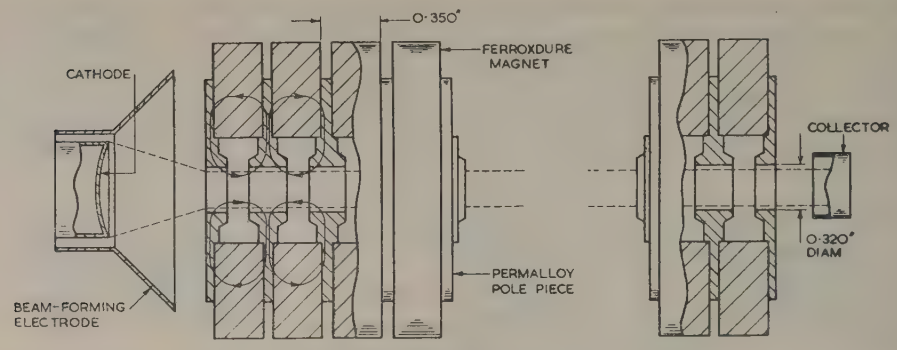


Fig. 28.—Production of long focused electron beams.

The electron beam is injected abruptly into the magnetic field of about 1 000 oersteds. To save weight, a periodic magnetic field is employed.

the tunnel has length l_t and radius r_t, the beam perveance has a maximum value of

$$\mathscr{P}_{max} = 9{\cdot}75 \times 10^{-6}(r_t/l_t)^2 \quad . \quad . \quad . \quad (82)$$

The current density at the minimum cross-section is greater than that at the ends of the cylinder by a factor of 5·66.

For a long electron beam, as required, say, in a travelling-wave tube, $r_t \ll l_t$ and the limiting perveance is low. Positive-ion focusing is possible but has not found favour.[592] In practice, beams of high perveance are invariably obtained by magnetic focusing. In confined focusing,[492] a cathode of substantially the required beam diameter is immersed in an axial magnetic field which is made strong enough to restrict the transverse motion to an acceptable amount. In a more refined version, the electron paths in the gun region are designed to lie along the lines of magnetic field.

In Brillouin[399] flow, the electrons, rotating about the beam axis, experience an inward force which just balances the outward one due to space-charge repulsion and centrifugal force. The beam radius, r_b, is obtained from the relation

$$I_0 = \frac{\pi}{\sqrt{2}}\epsilon_0\left(\frac{e}{m}\right)^{3/2} B^2 r_b^2 \left(V_0 - \frac{eB^2r_b^2}{8m}\right)^{1/2} \quad . \quad . \quad (83)$$

The maximum perveance becomes

$$\mathscr{P}_{max} = 25{\cdot}4 \times 10^{-6} \quad . \quad . \quad . \quad . \quad . \quad (84)$$

the condition being such that the electrons all travel with the same axial velocity. This velocity corresponds to the potential, $\frac{1}{3}V_0$, on the axis of the beam, the remaining two-thirds of the total beam energy being rotational. In practice, much less intense beams, with $\mathscr{P} = 1\text{–}2 \times 10^{-6}$, are preferred. For example, a beam of 200 mA at 3 kV can be projected down a tunnel of radius 1·4 mm and length 33 cm with a loss to the wall of only 1 mA. A typical method[608] of achieving Brillouin flow is shown in Fig. 28.

Focusing magnetic fields may be produced by solenoids for experimental purposes or by permanent magnets for fixed operation. A field of 500 oersteds over a distance of 5¾ in can, for example, be given[460] by a pair of permanent magnets weighing 3 lb. Eqn. (83) shows that the focusing effect depends upon H^2, a fact which has given rise to periodic[470, 555] or 'strong' focusing. At every plane where the magnetic field changes sign, the direction of electron rotation reverses but otherwise the beam is unaltered. The use of an array of N short magnets means that the stray magnetic field is considerably reduced,[527, 528] the gain factor lying between N and N^2. Several practical assemblies have been developed[403, 408, 409, 565] and one arrangement, emp ing a magnet of weight 1 lb 5 oz, is shown in Fig. 28.

Periodic electrostatic focusing[596] has not found much prac application[381] because the beam perveances are lower. Ce fugal methods have been proposed[410, 411, 413, 472] in which beam is set spinning at the gun region before entering the re of periodic electrostatic focusing fields. Unless a spi travelling circuit-wave or a collector which 'unwinds' the b is employed, the angular beam energy in these centrif systems is wasted. In another type of periodic electros focusing—slalom flow[418, 419, 420]—the electrons follow an potential in winding their way in sinuous fashion throug array of conductors.

Part 3. NOISE IN ELECTRON TUBES

(8) OSCILLATORS

The output signal from an electron tube is never a pe periodic wave: it is always slightly modified with ran fluctuations. This noise may be a function of the electr processes inside the tube, as well as the properties of structures associated with the microwave circuit. The output usually consists of a carrier, wide-band background and narrow-band modulation sidebands. The general sourc noise in electron tubes are well known:[676] they include noise,[681, 682] partition noise[633] and the modification du transit time.[657, 698]

Noise phenomena are important even when the tub employed as a generator. Wide-band noise in the local oscil output of a superheterodyne receiver beats with the carri produce sidebands at the intermediate frequency. For a ty reflex klystron, the 723 A/B, the power density of noi 2×10^{-12} watt per Mc/s at 30 Mc/s separation from the ca Since this tube generates about 50 mW of power, the noi this 1 Mc/s band is approximately 104 dB below the ca level. If the mixer of the receiver requires 1·5 mW of oscillator power, such a tube would also supply 6×10^{-14} per Mc/s of noise.

The performance of electron tubes as transmitters is influe by random fluctuations near the carrier frequency. These be considered to be made up of simultaneously existing pure and pure f.m. sidebands, with full correlation if both are ca by the same event, and zero correlation if they are not. usually f.m. noise which is of practical importance an causes have been studied.[688] Other theoretical and experim studies of noise have been concerned with klystrons,[635, 648, 6 magnetrons[638, 664, 666] and M-type carcinotrons.[643]

(9) AMPLIFIERS

It is as low-level amplifiers that the noise characteristics of ectron tubes have received most attention. Grid-controlled bes[672] may be used as such amplifiers, but their noise increases th operating frequency. For example, the GL416B triode s noise temperatures of 120° K and 1 200° K at 100 Mc/s and Gc/s respectively. Klystrons show excess noise due to parti- n effects and conversion of thermal emission velocities to nsity fluctuations, and are rarely used as low-level amplifiers. Good low-level performance is achieved with beam-type tubes d, with care, only noise which is due to fluctuations in current nsity and emission velocity predominates. The fluctuations esent at the input end of the electron beam excite space-charge ves and several theories treating the beam as a wave transmis- n system have been given.[637, 652, 680, 682] More accurate sults are obtained[634, 681, 685, 686, 697] by taking into account the noothing of the current fluctuations near the potential minimum. It was predicted[673, 674] that the cathode noise varies periodically ng the beam. If I_{min} and I_{max} are the noise currents, then

$$|I_{min}||I_{max}| = A\frac{\omega}{\omega_p}\frac{kT_c}{eV_0}(ef_bI_0) \quad . \quad . \quad . \quad (85)$$

ere A is a constant near unity. The noise pattern is made of two uncorrelated standing-wave patterns, due respectively the current and velocity fluctuations. If I_1 and I_2 are their axima, the minima being zero, then

$$|I_{min}||I_{max}| = |I_1||I_2| \sin\psi \quad . \quad . \quad . \quad (86)$$

ere ψ is the spatial phase angle. This analysis is supported by e experimental results[647] shown in Fig. 29(*a*).

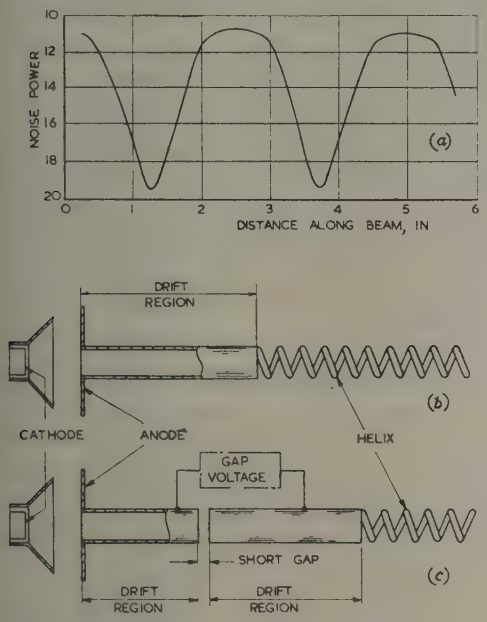


Fig. 29.—Noise on an electron beam.
(*a*) Noise amplitude as a function of distance.
(*b*) Noise reduction by optimum-length drift-tube.
(*c*) Further noise reduction with short velocity jump.

Lower tube noise can be secured by putting the input end of e slow-wave circuit at some critical position along the beam, shown in Fig. 29(*b*). A further improvement is obtained[694] by ploying drift tubes at different steady potentials. The result- g changes in impedance can, by correct choice of length and position, result in de-amplification of the noise. This method is illustrated in Fig. 29(*c*), the potential of the first drift tube being jumped to that of the second at the point where the noise current is a minimum. This second tube is adjusted so that there is also a critical distance between the jump and helix. Some care must be used in designing velocity jumps which, being produced by sharp potential discontinuities, behave like strong electron lenses. Theory and experiment predict[659] that such lenses tend to increase the noise in an electron beam as a consequence of the transformation of radial into axial velocity fluctuations.

Estimates of the theoretical minimum noise figure for beam-type tubes have been made.[651, 653, 654, 662, 679] If I_{min} were zero, the noise figure of the tube would be unity. A higher noise figure exists[636] as a consequence of the lack of correlation between the initial noise current and velocity fluctuations of the beam. Making certain assumptions about the velocity and current fluctuations at the potential minimum and employing lossless beam transducers, a noise figure of

$$F_{min} = 1 + (4 - \pi)^{1/2}T_c/T_0 \quad . \quad . \quad . \quad . \quad (87)$$

has been predicted.[675] For practical parameters, this gives 6 dB as the minimum noise figure for beam-type amplifiers. A rather better value is obtained by considering the conditions at the potential minimum. One analysis[696] gives results reasonably correct at the lower microwave frequencies, while computations[691, 692] for a particular geometry give a noise figure which varies with frequency.

A calculation procedure for low-noise travelling-wave tubes has shown[639] that greater design flexibility is achieved by the use of three or more separate anodes. In such tubes, proper design of the electron gun is desirable,[661] with uniform emission[634] from the cathode. Tube noise may be affected by thermal-velocity spread,[668] emission of secondary electrons,[671] the focusing magnetic field[677] and partition effects.[678] Recent work[687] has shown that an electron beam may be 'cooled' by certain processes.

Information on the design of low-noise tubes has been given[655, 658, 667] and measurements of their properties reported.[631, 690] These techniques have been incorporated in the construction of several practical tubes.[650, 656, 669, 684] Examples at 3 Gc/s have given[642, 670] noise figures of 8 dB, a typical gain being 16 dB at a bandwidth of 100 Mc/s and beam current of 400 μA. More recently,[641] a noise figure less than 4 dB has been achieved. Low-noise tubes have also been developed for 4 Gc/s[632, 640] and 9 Gc/s.[695]

The conditions have been studied for low-noise operation of transverse-field,[693] double-stream[665] and backward-wave[649] amplifiers. A general method of design for the last group, utilizing the circle diagram of impedance, has been presented:[645] it was shown that low-voltage tubes provide optimum noise-figure and tuning-range characteristics, conclusions which are verified in practice.[644, 646]

(10) ACKNOWLEDGMENTS

The author is indebted to Dr. T. S. England for useful discussion on this subject and to Mr. I. A. D. Lewis and Dr. H. W. Duckworth for reading the manuscript.

Part 4. BIBLIOGRAPHY

Part 1

(1) Adler, R., Kromhout, O. M., and Clavier, P. A.: 'Resonant Behaviour of Electron Beams in Periodically Focused Tubes for Transverse Signal Fields', *Proceedings of the Institute of Radio Engineers*, 1955, 43, p. 339.
(2) Alfven, H.: 'A High Frequency Tube Based on a New Type of Electron Motion', *L'Onde Electrique*, 1957, 37, p. 65.
(3) Agdur, B. N.: 'Experimental Observation of Double-Stream Amplification', *Acta Polytechnica*, 1951, No. 86, p. 9.

(4) AGDUR, B. N.: 'On the Interaction between Microwave Fields and Electrons, with Special Reference to the Strophotron', *Ericsson Technics*, 1957, **13**, p. 3.
(5) ANDREWS, C. L.: 'Microwave Triode Oscillators', *Review of Scientific Instruments*, 1957, **28**, p. 443.
(6) ASHLEY, D. E. T. F., COCKHILL, T. D., HASSELL, A. F., and JENKINS, R. O.: 'Interaction within the Attenuator of a High-Power T.W.T.', *Journal of Electronics and Control*, 1958, **5**, p. 62.
(7) BARFORD, N. C., and BOWMAN-MANIFOLD, M.: 'Elementary Theory of Velocity-Modulation Oscillators', *Journal I.E.E.*, 1947, **94**, Part III, p. 302.
(8) BARNES, C. W.: 'Power and Gain Limitations of Helix-Type Travelling-Wave Tubes', *Le Vide*, 1957, **12**, p. 43.
(9) BEAM, W. R.: 'On the Possibility of Amplification in Space-Charge-Potential-Depressed Electron Streams', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 454.
(10) BECK, A. H. W.: 'Space Charge Waves' (Pergamon, London, 1958).
(11) BECK, A. H. W.: 'Velocity-Modulated Thermionic Tubes' (University Press, Cambridge, 1948).
(12) BEGGS, J. E., and LAVOO, N. T.: 'A Triode Useful to 10,000 Mc/s', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 15.
(13) BELL, J.: 'Triodes for Very Short Waves', *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 187.
(14) BELL, J., GAVIN, M. R., JAMES, E. G., and WARREN, G. W.: 'Triodes for Very Short Waves', *ibid.*, 1946, **93**, Part III, p. 833.
(15) BELL, R. L., and HILLIER, M.: 'An 8 mm Klystron Power Oscillator', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 1155.
(16) BENNETOT, M. DE: 'Experimental Investigation of the Oscillation Mode and Bunching Efficiency of a Reflex Klystron', *Annales de Radioélectricité*, 1953, **8**, p. 103.
(17) BENNETOT, M. DE: 'Experimental Investigation of the Operation of the Reflection Space of a Reflex Klystron', *ibid.*, 1954, **9**, p. 150.
(18) BENNETT, W. P.: 'New Beam Power Tubes for UHF Service', *Transactions of the Institute of Radio Engineers*, 1956, **ED–3**, No. 1, p. 57.
(19) BENNETT, W. P., and KAZANOWSKI, H. F.: 'One Kilowatt Tetrode for UHF Transmitters', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 13.
(20) BERNIER, J., and LEBOUTET, H.: 'Possibility of obtaining Very Short Continuous Waves, using a Reflex Klystron Supplying Energy at High-Order Harmonics of the Fundamental Oscillation', *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences (Paris)*, 1954, **239**, p. 796.
(21) BERTEROTTIERE, and CONVERT, G.: 'On Certain Effects of the Space Charge in Travelling-Wave Valves', *Annales de Radioélectricité*, 1950, **5**, p. 168.
(22) BIRDSALL, C. K., and WHINNERY, J. R.: 'Waves in an Electron Stream with General Admittance Walls', *Journal of Applied Physics*, 1953, **24**, p. 314.
(23) BIRDSALL, C. K., BREWER, G. R., and HAEFF, A. V.: 'The Resistive-Wall Amplifier', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 865.
(24) BIRDSALL, C. K.: 'Rippled Wall and Rippled Stream Amplifiers', *ibid.*, 1954, **42**, p. 1628.
(25) BIRDSALL, C. K., and BREWER, G. R.: 'Traveling-Wave-Tube Characteristics for Finite Values of C', *Transactions of the Institute of Radio Engineers*, 1954, **ED–1**, p. 1.
(26) BLEANEY, B.: 'Electronic Tuning of Reflection Klystrons', *Wireless Engineer*, 1948, **25**, p. 6.
(27) BLIOKH, P. V.: 'High-Frequency Oscillations in Electron Beams with a Periodically Varying Velocity', *Radiotekhnika i Elektronika*, 1957, **2**, No. 1, p. 92.
(28) BLIOKH, P. V., and FAINBERG, Y. B.: 'Space Charge Waves in Velocity-Modulation Electron Beams', *Zhurnal Tekhnicheskoi Fiziki*, 1956, **26**, p. 530.
(29) BLOOM, S., and PETER, R. W.: 'Transmission-Line Analog of a Modulated Electron Beam', *RCA Review*, 1954, **15**, p. 95.
(30) BOBENRIETH, A.: 'A Travelling-Wave Valve for 4 cm Wavelength', *Le Vide*, 1956, **11**, p. 290, and *L'Onde Électrique*, 1957, **37**, p. 86.
(31) BOBENRIETH, A., and CAHEN, O.: 'Travelling-Wave Valves for 4 cm Waves: Research and Development at the Centre National d'Études des Télécommunications', *L'Onde Électrique*, 1956, **36**, p. 307.
(32) BODE, H. W.: 'Network Analysis and Feedback Amplifier Design' (Van Nostrand, New York, 1945).
(33) BOHLKE, P. G., and BREEDEN, F. G.: 'External Cavity Klystron', *Electronics*, July, 1947, p. 114.
(34) BOONE, E. M., UENOHARA, M., and DAVIS, D. T.: 'A Barkhausen–Kurz Oscillator at Centimetre Wavelengths', *Transactions of the Institute of Radio Engineers*, 1958, **ED–5**, p. 196.
(35) BOWEN, A. E., and MUMFORD, W. W.: 'A New Microwave Triode: its Performance as a Modulator and as an Amplifier', *Bell System Technical Journal*, 1950, **29**, p. 531.
(36) BRANCH, G. M., and MIHRAN, T. G.: 'Plasma-Frequency Reduction Factors in Electron Beams', *Transactions of the Institute of Radio Engineers*, 1955, **ED–2**, No. 2, p. 3.
(37) BRESLER, A. D., JOSHI, G. H., and MARCUVITZ, N.: Orthogonality Properties for Modes in Passive and Active Uniform Waveguides', *Journal of Applied Physics*, 1958, **29**, p. 794.
(38) BREWER, G. R., and BIRDSALL, C. K.: 'Traveling-Wave Tube Propagation Constants', *Transactions of the Institute of Radio Engineers*, 1957, **ED–4**, No. 2, p. 140.
(39) BRIDGES, T. J., and CURNOW, H. J.: 'Experimental 8-mm Klystron Power Amplifiers', *Proceedings of the Institute of Radio Engineers*, 1958, **46**, p. 430.
(40) BRILLOUIN, L.: 'The Travelling-Wave Tube (Discussion of Waves for Large Amplitudes)', *Journal of Applied Physics*, 1949, **20**, p. 1196.
(41) BROADWAY, L. F.: 'Velocity-Modulation Valves', *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 182.
(42) BROADWAY, L. F., MILNER, C. J., PETRIE, D. R., SCOTT, W. J., and WRIGHT, G. P.: 'Velocity-Modulation Valves', *ibid.*, 1946, **93**, Part IIIA, p. 855.
(43) BROWN, W. C.: 'Platinotron Increases Search Radar Range', *Electronics*, August, 1957, p. 164.
(44) BRUCK, G. G.: 'Modified Cavity Oscillator for the Generation of Microwaves', *Journal of Applied Physics*, 1947, **18**, p. 843.
(45) BRUCK, L.: 'Comparison of the Measured Valves for the Linear Gain of the Travelling Wave Valve with the Values indicated by Various Theories', *Annales de Radioélectricité*, 1949, **4**, p. 222.
(46) BRUCK, L.: 'The Travelling-Wave Valve', *Archiv für Elektrotechnik*, 1950, **39**, p. 633.
(47) BRUYSTEN, J.: 'Graphical Determination of Reflex-Klystron Characteristics', *Philips Research Reports*, 1955, **10**, p. 81.
(48) CACHERIS, J. C., JONES, G., and DIEHL, L.: 'Magnetic Tuning of Klystron Cavities', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 1017.
(49) CALDWELL, J. J.: 'High Power Traveling-Wave Tube Gain and Saturation Characteristics as a Function of Attenuator Configuration and Resistivity', *Transactions of the Institute of Radio Engineers*, 1953, **ED–4**, No. 5, p. 28.
(50) CALDWELL, J. J., and HOCK, O. L.: 'Large Signal Behaviour of High Power Traveling-Wave Amplifiers', *ibid.*, 1956, **ED–3**, No. 1, p. 6.
(51) CARTER, C. J., and CORNETET, W. H.: 'A Low Voltage One Centimetre Retarding-Field Oscillator', *ibid.*, 1956, **ED–3**, No. 3, p. 124.
(52) CARTER, C. J., CORNETET, W. H., and THURSTON, M. O.: 'Operation Application of the Retarding-Field Oscillator at Millimetre Wavelen *Le Vide*, 1956, **11**, p. 281, and *L'Onde Électrique*, 1957, **37**, p. 65.
(53) CARTER, C. J., and CORNETET, W. H.: 'Low-Voltage Operation of the Retar Field Oscillator at X-Band and in the Millimetre Wavelength Re *Transactions of the Institute of Radio Engineers*, 1958, ED–5, No. 3, p. 1
(54) CHALBERG, H. W. A.: 'R.F. Performance of a UHF Triode', *Proceedings Institute of Radio Engineers*, 1953, **41**, p. 46.
(55) CHALK, G. O., MANLEY, B. W., and NORRIS, V. J.: 'A Five-Cavity X- Klystron Amplifier', *Journal of Electronics*, 1956, **2**, p. 50.
(56) CHODOROW, M., and FAN, S. P.: 'A Floating-Drift-Tube Klystron', *Proce of the Institute of Radio Engineers*, 1953, **41**, p. 25.
(57) CHODOROW, M., GINZTON, E. L., and NALOS, E. J.: 'Debunching of Ele Beams constrained by Strong Magnetic Fields', *ibid.*, 1953, **41**, p. 999.
(58) CHODOROW, M., GINZTON, E. L., NEILSON, I. R., and SONKIN, S.: 'Desig Performance of a High-Power Pulsed Klystron', *ibid.*, 1953, **41**, p. 1584.
(59) CHODOROW, M., and NALOS, E. J.: 'The Design of High-Power Traveling- Tubes', *ibid.*, 1956, **44**, p. 649.
(60) CHODOROW, M., and WESTBURG, V. B.: 'Space-Charge Effects in Reflex strons', *ibid.*, 1951, **39**, p. 1548.
(61) CHRISTENSEN, W. V., and WATKINS, D. A.: 'Helix Millimetre-Wave T *ibid.*, 1955, **43**, p. 93.
(62) CHU, L. J., and JACKSON, J. D.: 'Field Theory of Traveling-Wave T *ibid.*, 1948, **36**, p. 853.
(63) CLARKE, G. M.: 'A Symbolism for Microwave-Valve Classification', *Proce I.E.E.*, Monograph No. 14 R, November, 1951 (**99**, Part IV, p. 24).
(64) CLARRICOATS, P. J. B., HAYES, A. G., and HARVEY, A. F.: 'A Survey Theory and Applications of Ferrites at Microwave Frequencies', *ibid.*, No. 2229 R, October, 1956 (**104 B**, Suppl. No. 6, p. 267).
(65) COCKROFT, H. S., and PICKIN, J. R.: 'X-Band Klystrons for High Power Operation', *Journal of Electronics*, 1956, **1**, p. 359.
(66) COHEN, S. A.: 'Traveling-Wave Tube Gain Fluctuations with Freque *Transactions of the Institute of Radio Engineers*, 1957, **ED–4**, No. 1, p. 7
(67) COOK, J. S., KOMFNER, R., and SUHL, H.: 'Non-reciprocal loss in Trav Wave Tubes using Ferrite Attenuators', *Proceedings of the Institute of Engineers*, 1954, **42**, p. 1188.
(68) COULSON, R. B., and ROBINSON, F. N. H.: 'A Medium-Power Travelling- Tube for 2000 Mc/s', *Marconi Review*, 1955, **18**, p. 48.
(69) CRAVEN, G.: 'Disc-Seal Triode Amplifiers', *Wireless Engineer*, 1956, p. 179.
(70) CRUMLY, C. B.: 'A UHF Traveling Amplifier Wave Tube employing an Ele statically Focused Hollow Beam', *Transactions of the Institute of Engineers*, 1956, **ED–3**, No. 1, p. 62.
(71) CURRIE, M. R., and WHINNERY, J. R.: 'The Cascade Backward Wave Ampl a High Gain Voltage-Tuned Filter for Microwaves', *Proceedings of the Ins of Radio Engineers*, 1955, **43**, p. 1617.
(72) CURRIE, M. R., and FORSTER, D. C.: 'The Gain and Bandwidth Character of Backward-Wave Amplifiers', *Transactions of the Institute of Radio Engi* 1957, **ED–4**, No. 1, p. 24.
(73) CUTLER, C. C.: 'The Calculation of Traveling-Wave-Tube Gain', *Proce of the Institute of Radio Engineers*, 1951, **39**, p. 914.
(74) CUTLER, C. C., and BRANGACCIO, D. J.: 'Factors affecting Traveling-Wave- Power Capacity', *Transactions of the Institute of Radio Engineers*, 1953, E No. 3, p. 9.
(75) CUTLER, C. C.: 'The Nature of Power Saturation in Traveling-Wave Tu *Bell System Technical Journal*, 1956, **35**, p. 841.
(76) DAIN, J.: 'Factors in the Design of Power Amplifiers for Ultra-High quencies', *Journal of Electronics*, 1955, **1**, p. 35.
(77) DAVIS, J. I.: 'Technique of Pulsing Low Power Reflex Klystrons', *Transac of the Institute of Radio Engineers*, 1956, **MTT–4**, p. 40.
(78) DENCH, E.: 'A Voltage-tuned High-power Microwave Oscillator, *ceedings of the National Electronics Conference (Chicago)*, 1954, **9**, p.
(79) DENIS, M.: 'Generalities on Travelling-Wave-Valve Feedback Self-Oscilla Theory of the Reflex Travelling-Wave Valve', *Annales de Radioélectr* 1952, **7**, p. 169.
(80) DENMAN, E. D., and LALLY, P. M.: 'High Gain Traveling-Wave Amplif *Transactions of the Institute of Radio Engineers*, 1957, **ED–3**, p. 83.
(81) DERFLER, H.: 'An Electromagnetic Difference-Equation of Importance i Theory of Travelling-Wave Tubes', *Zeitschrift für angewandte Mather und Physik*, 1955, **6**, p. 104.
(82) DIEMER, G., and KNOL, K. S.: 'Low-level Triode Amplifier for Microwa *Philips Research Reports*, 1950, **5**, p. 153.
(83) DIEMER, G., and RODENHUIS, K.: 'Optimum Geometry of Microwave Amp Valves', *ibid.*, 1952, **7**, p. 36.
(84) DODDS, W. J., PETER, R. W., and KAISEL, S. F.: 'New Development in Trave Wave Tubes', *Electronics*, February, 1953, p. 130.
(85) DOEHLER, O., KLEEN, W., and PALLUEL, P.: 'Travelling-Wave Valves as lators with Wideband Electronic-Tuning', *Annales de Radioélectricité*, **4**, p. 68.
(86) DOEHLER, O., and KLEEN, W.: 'The Efficiency of the Travelling-Wave V *Archiv der Elektrischen Übertragung*, 1950, **4**, p. 207.
(87) DOEHLER, O.: 'The Control of Electron Streams by means of Progressive W (Travelling-Wave Tubes)', *Funk und Ton*, 1950, **4**, p. 493.
(88) DOEHLER, O., and KLEEN, W.: 'Non-Linear Phenomena in Travelling Tubes', *Annales de Radioélectricité*, 1948, **3**, p. 124.
(89) DOEHLER, O., EPSZTEIN, B., and ARNAUD, J.: 'New Circuits for T.W.T.', *L' Électrique*, 1957, **37**, p. 86.
(90) DOOLITTLE, H. D.: 'The Design and Life of Planar Microwave Transmi Tubes', *Transactions of the Institute of Radio Engineers*, 1957, VC–8, p.
(91) DORE, B. V.: 'Velocity-Jump Amplification at 10,000 Mc/s', *Canadian Jour Physics*, 1957, **35**, p. 742.
(92) DORGELO, E. G.: 'Some Technological Aspects of UHF Triode Design', *Le* 1957, **12**, p. 3.
(93) DUNN, D. A.: 'Traveling-Wave Amplifiers and Backward-Wave Oscillato VHF', *Transactions of the Institute of Radio Engineers*, 1957, **ED–4**, N p. 246.
(94) DUNN, D. A., and HARMAN, W. A.: 'An Experimental Transverse-Cu Traveling-Wave Tube', *Proceedings of the Institute of Radio Engineers*, **44**, p. 888.
(95) DUNN, D. A., HARMAN, W. A., FIELD, L. M., and KINO, G. S.: 'Theory Transverse-Current Traveling-Wave Tube', *ibid.*, 1956, **44**, p. 879.
(96) DUNN, D. A.: 'Understanding the Backward-Wave Oscillator', *Elect Industries*, 1958, **17**, p. 72.
(97) EBERS, J. J.: 'Retarding-field Oscillators', *Proceedings of the Institute of Engineers*, 1952, **40**, p. 138.
(98) EICHIN, W., MEYERER, P., VEITH, W., and ZINKE, O.: 'High-Gain Trave Wave Output Valve with Permanent Magnet', *Fernmeldetechnische Zeits* 1955, **8**, p. 369.
(99) FANK, F. B., and WADE, G.: 'Traveling-Wave Tube Limiters', *Transactio the Institute of Radio Engineers*, 1957, **ED–4**, No. 2, p. 148.

) FEENBERG, E.: 'Elementary Treatment of Longitudinal Debunching on a Velocity Modulation System', *Journal of Applied Physics*, 1946, 17, p. 852.

) FIELD, L. M.: 'Recent Developments in Traveling Wave Valves' *Electronics*, January, 1950, p. 100.

) FIELD, L. M., TIEN, P. K., and WATKINS, D. A.: 'Amplification by Acceleration and Deceleration of a Single Velocity Stream', *Proceedings of the Institute of Radio Engineers*, 1951, 39, p. 194.

) FIELD, L. M.: 'Some Slow-Wave Structures for Traveling-Wave Tubes', *ibid.*, 1949, 37, p. 34.

) FLETCHER, R. C.: 'Helix Parameters in Traveling-Wave Tube Theory', *ibid.*, 1950, 38, p. 413.

) FOSTER, J.: 'Grounded-Grid Amplifier for Very Short-Waves', *Journal I.E.E.*, 1946, 93, Part IIIA, p. 868.

) FREMLIN, J. H., GENT, A. W., PETRIE, D. P. R., WALLIS, P. J., and TOMLIN, S. G.: 'Principles of Velocity Modulation', *ibid.*, 1946, 93, Part IIIA, p. 875.

) FREMLIN, J. H.: 'Velocity Modulation:—Parts 1 and 2', *Electronic Engineering*, 1947, 19, pp. 147 and 199.

) FREY, W., and LUDI, F.: 'Theory of the Travelling-Wave Valve', *Zeitschrift für angewandte Mathematik und Physik*, 1950, 1, p. 237.

) FRIEDMAN, B.: 'Amplification of the Travelling-Wave Tube', *Journal of Applied Physics*, 1951, 22, p. 443.

) GABOR, D.: 'Plasma Oscillations', *British Journal of Applied Physics*, 1951, 2, p. 209.

) GENT, A. W.: 'Small-Amplitude Theory and Starting Current', *Journal I.E.E.*, 1946, Part IIIA, p. 883.

) GEPPERT, D. V.: 'Analysis of Traveling-Wave Tubes with Tapered Velocity Parameter', *Proceedings of the Institute of Radio Engineers*, 1958, 46, p. 1658.

) GIACOLETTO, L. J., and JOHNSON, H.: 'UHF Triode Design in Terms of Operating Parameters and Electrode Spacings', *ibid.*, 1953, 41, p. 51.

) GINZTON, E. L., and HARRISON, A. E.: 'Reflex-Klystron Oscillators', *ibid.*, 1946, 34, p. 97.

) GITTINS, J. F., ROCK, N. H., and SULLIVAN, A. B.: 'An Experimental High Power Pulsed Travelling Wave Tube', *Journal of Electronics*, 1957, 3, p. 267.

) GODFREY, R. M., HUMPHRIES, B. L., ALLIN, P. E. V., and MOTT, G.: 'Applications of Ferrites at 3 cm Wavelength', *Proceedings I.E.E.*, Paper No. 2197 R, October, 1956 (104 B, Suppl. No. 6, p. 355).

) GOUDET, G.: 'Recent Developments in Amplifying Valves for Centimetre Waves', *Annales des Télécommunications*, 1948, 3, p. 445.

) GOUDET, G.: 'The Amplification of Centimetre Waves', *L'Onde Électrique*, 1950, 30, p. 8.

) GOULD, R. W.: 'A Coupled Mode Description of the Backward-Wave Oscillation and the Kompfner Dip Condition', *Transactions of the Institute of Radio Engineers*, 1955, ED–2, No. 4, p. 37.

) GOULD, R. W.: 'Characteristics of Traveling-Wave Tubes with Periodic Circuits', *ibid.*, 1958, ED–5, p. 186.

) GRIFFITHS, J. H. E.: 'The Development of Radio Valves', *Journal I.E.E.*, 1946, 93, Part IIIA, p. 173.

) GROW, R. W., DUNN, D. A., MCLAUGHLIN, J. W., and LAGERSTROM, R. P.: 'A 20 to 40 kMc/s Backward-Wave Oscillator', *Transactions of the Institute of Radio Engineers*, 1958, ED–5, No. 3, p. 152.

) GROW, R. W., and WATKINS, D. A.: 'Backward-Wave Oscillator Efficiency', *Proceedings of the Institute of Radio Engineers*, 1955, 43, p. 848.

) GUENARD, P., WARNECKE, R., and FAUVE, C.: 'On the Efficiency of Velocity Modulation Valves', *Annales de Radioélectricité*, 1948, 3, p. 302, 1949, 4, p. 92.

) GUENARD, P., DOEHLER, O., EPSZTEIN, B., and WARNECKE, R.: 'New VHF Valves with Wide Electronic Tuning Band', *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* (*Paris*), 1952, 235, p. 236.

) GUENARD, P., DOEHLER, O., and WARNECKE, R.: 'Properties of Lines with Periodic Structures', *ibid.*, 1952, 235, p. 32.

) GUREWITSCH, A. M., and WHINNERY, J. R.: 'Microwave Oscillators using Disc-Seal Tubes', *Proceedings of the Institute of Radio Engineers*, 1947, 35, p. 462.

) HAEFF, A. V.: 'Space-Charge Wave Amplification Effects', *Physical Review*, 1948, 74, p. 1532.

) HAEFF, A. V.: 'The Electron-Wave Tube—A Novel Method of Generation and Amplification of Microwave Energy', *Proceedings of the Institute of Radio Engineers*, 1949, 37, p. 4.

) HAGGBLOM, H., and TOMNER, S.: 'Developments of the Strophotron', *Ericsson Technics*, 1956, 12, p. 165.

) HAHN, W. C.: 'Small Signal Theory of Velocity Modulated Electron Beams', *General Electric Review*, 1939, 42, p. 258.

) HAMILTON, D. R., KNIPP, J. K., and KUPER, J. B. H.: 'Klystrons and Microwave Triodes' (McGraw-Hill, New York, 1948).

) HARMAN, W. W.: 'Fundamentals of Electronic Motion' (McGraw-Hill, New York, 1953).

) HARRIES, J. H. OWEN: 'Cavity Resonators and Electron Beams', *Wireless Engineer*, 1947, 24, p. 135.

) HARRIS, L. A.: 'The Effect of an Initial Velocity Spread on Klystron Performance', *Transactions of the Institute of Radio Engineers*, 1958, ED–5, No. 3, p. 157.

) HARRIS, W. A., and THOMPSON, J. J.: 'A UHF-VHF Tuner using Pencil Tubes', *RCA Review*, 1955, 16, p. 281.

) HARRISON, A. E.: 'Klystron Tubes' (McGraw-Hill, New York, 1947).

) HARVEY, A. F.: 'High Frequency Thermionic Tubes' (Chapman and Hall, London, 1942).

) HAUS, H. A., and BOBROFF, D. L.: 'Small-Signal Power Theorem for Electron Beams', *Journal of Applied Physics*, 1957, 28, p. 694.

) HECHTEL, R.: 'Modern Reflex Klystrons', *Archiv der Elektrischen Übertragung*, 1956, 10, p. 133.

) HEFFNER, H.: 'Backward-Wave Tube', *Electronics*, October, 1953, p. 135.

) HEFFNER, H.: 'Analysis of the Backward-Wave Traveling-Wave Tube', *Proceedings of the Institute of Radio Engineers*, 1954, 42, p. 930.

) HEIL, O., and EBERS, J. J.: 'A New Wide Range High Frequency Oscillator', *ibid.*, 1950, 38, p. 645.

) HELMKE, G. E.: 'A Hairpin-Tube Backward-Wave Oscillator', *Bell Laboratories Record*, 1953, 31, p. 286.

) HEMPSTEAD, C. F., and STRAND, A. R.: 'A Versatile Source of Millimetre Waves', *ibid.*, 1957, 35, p. 241.

) HOLLENBERG, A. V.: 'Experimental Observation of Amplification by Interaction between Two Electron Streams', *Bell System Technical Journal*, 1949, 28, p. 52.

) HOLLMAN, H. E.: 'Barkhausen Electron Oscillations as the Basis of Velocity-Modulated Valves', *Hochfrequenztechnik und Elektroakustic*, 1957, 65, p. 112.

) HOLMBOE, L. W., and ETTENBERG, M.: 'Development of a Medium Power L-Band Traveling-Wave Amplifier', *Transactions of the Institute of Radio Engineers*, 1957, ED–4, p. 78.

) HORTON, C. E.: 'Development of a UHF Grounded Grid Amplifier', *Proceedings of the Institute of Radio Engineers*, 1953, 41, p. 73.

) HULSTER, F.: 'Transmitter Valves with Control Grid', *Le Vide*, 1956, 11, No. 65, p. 226, and *L'Onde Électrique*, 1957, 37, p. 86.

) HUTTER, R. G. E.: 'Space-Charge-Wave Amplifier Tubes, Basic Principles of Operation', *Sylvania Technologist*, 1952, 5, p. 94, and 1953, 6, p. 6.

(152) HUTTER, R. G. E., CUTLER, S., and GREENBERG, H.: 'Traveling-Wave Tubes', *Radio and Television News*, 1954, 51, p. 23.

(153) HUTTER, R. G. E.: 'Travelling-Wave Tubes', Advances in Electronics and Electron Physics, Vol. VI (Academic Press, New York, 1954).

(154) VAN IPEREN, B. B.: 'Reflex Klystrons for Millimetre Waves', *Le Vide*, 1956, 11, p. 264, and *L'Onde Électrique*, 1957, 37, p. 86.

(155) JAMIESON, H. W., and WHINNERY, J. R.: 'Study of Transit Time Effects in Disc-Seal Power-Amplifier Triodes', *Proceedings of the Institute of Radio Engineers*, 1948, 36, p. 76.

(156) JASBERG, J. H.: 'Improvement of Power Output from Pulsed Klystrons', *ibid.*, 1954, 42, p. 849.

(157) JOHNSON, H. R., and WHINNERY, J. R.: 'Traveling Wave Oscillator Tunes Electronically', *Electronics*, August, 1953, p. 177.

(158) JOHNSON, H. R., and WEGLEIN, R. D.: 'Backward-Wave Oscillators for 8 000–18 000 Mc/s Band', *Transactions of the Institute of Radio Engineers*, 1957, ED–4, p. 180.

(159) JOHNSON, H. R.: 'Backward Wave Oscillators', *Proceedings of the Institute of Radio Engineers*, 1955, 43, p. 684.

(160) JONES, E. M. T.: 'An Internal-Feedback Traveling-Wave-Tube Oscillator', *ibid.*, 1952, 40, p. 478.

(161) JONES, M. C.: 'Ground-Grid R.F. Voltage Amplifiers', *ibid.*, 1944, 32, p. 423.

(162) KARP, A.: 'Backward-Wave Oscillator Experiments at 100 to 200 kMc/s', *ibid.*, 1957, 45, p. 496.

(163) KARP, A.: 'Traveling-Wave Tube Experiments at Millimetre Wavelengths with a New, Easily Built, Space Harmonic Circuit', *ibid.*, 1955, 43, p. 41.

(164) KAUFMAN, I.: 'High-Energy Bunched Beam Analyser', *Journal of Applied Physics*, 1953, 24, p. 1413.

(165) KEIL, A., and PARZEN, P.: 'Nonlinear Wave Propagation in Traveling-Wave Amplifiers', *Transactions of the Institute of Radio Engineers*, 1955, ED–2, No. 4, p. 26.

(166) KENT, G.: 'Space Charge Waves in Inhomogeneous Electron Beams', *Journal of Applied Physics*, 1954, 25, p. 32.

(167) KHOKHLOV, R. V.: 'Theory of Self-Oscillation in a Reflex Klystron', *Zhurnal Tekhnicheskoi Fiziki*, 1955, 25, p. 2492.

(168) KIRYUSHIN, V. P.: 'Experimental Investigation of a Backward-Wave Valve with Bifilar Helix', *Radiotekhnika i Elektronika*, 1956, 1, p. 798.

(169) KLEEN, W. J.: 'Electronics of Microwave Tubes' (Academic Press, New York, 1958).

(170) KLEIN, W., and FRIZ, W.: 'A Twin-Helix Travelling-Wave Power Valve with 50 dB Gain at 4 kMc/s', *Nachrichtentechnische Zeitschrift*, 1956, 9, p. 476.

(171) KLEIN, W.: 'A Sensitive Method of Measuring the Reflections and Stability in High-Gain Travelling-Wave Tubes', *Archiv der Elektrischen Übertragung*, 1956, 10, p. 477.

(172) KLEIN, W., and FRIZ, W.: 'Characteristics of the Travelling-Wave Valve in the Dispersion Region of its Delay Line', *ibid.*, 1953, 7, p. 236.

(173) KLEIN, W., and POSCHL, K.: 'The Effect of Line Attenuation on the Power Gain of a Travelling-Wave Valve', *Fernmeldetechnische Zeitschrift*, 1953, 6, p. 509.

(174) KLEIN, W.: 'Measurement of the Most Important Parameters of Helix-Type Travelling-Wave Valves for Amplification and Efficiency Calculations', *Archiv der Elektrischen Übertragung*, 1954, 8, p. 404.

(175) KLEIN, W., and FRIZ, W.: 'The Performance of Travelling-Wave Valves at High Input Levels', *Fernmeldetechnische Zeitschrift*, 1954, 7, p. 349.

(176) KLEIN, W.: 'Delay Lines with Periodic Structures in Travelling-Wave Valves', *ibid.*, 1954, 7, p. 547.

(177) KLEIN, W.: 'History Classification and Physics of Very-High-Frequency Electron Valves', *Electrotechnische Zeitschrift*, 1955, 76A, p. 53.

(178) KLEIN, W. P. G.: 'Bihelical Traveling-Wave Tube with 50 dB Gain at 4 000 Mc/s', *Electrical Communication*, 1955, 32, p. 255.

(179) KLEINWACHTER, H.: 'Travelling-Wave Valves', *Elektrotechnik* (*Berlin*), 1950, 4, p. 245.

(180) KLEINWACHTER, H.: 'Travelling-Wave Valve with Sinuous Rectangular Waveguide', *Archiv der Elektrischen Übertragung*, 1952, 6, p. 460.

(181) KLINGER, H. H.: 'New Travelling-Wave Valves for Microwaves', *ibid.*, 1951, 5, p. 167.

(182) KOLESOV, L. N.: 'Energy Relations in a UHF Valve Oscillator', *Radiotekhnika*, 1956, 11, p. 27.

(183) KOIKE, Y., and YAMANAKA, S.: 'Amplification Constant for Microwave Triodes', *Technological Reports* (*Tohoku University*), 1951, 15, p. 14.

(184) KOLOMENSHI, A. A.: 'Radiation Emitted by a Uniformly Moving Electron in Electron Plasma in a Magnetic Field', *Comptes Rendus de l'Académie des Sciences de l'URSS*, 1956, 106, p. 982.

(185) KOMPFNER, R.: 'The Travelling-Wave Valve', *Wireless World*, 1946, 52, p. 369.

(186) KOMPFNER, R.: 'The Travelling-Wave Tube', *Wireless Engineer*, 1947, 24, p. 255.

(187) KOMPFNER, R.: 'The Traveling-Wave Tube as Amplifier at Microwaves', *Proceedings of the Institute of Radio Engineers*, 1947, 35, p. 124.

(188) KOMPFNER, R.: 'Travelling-Wave Tubes', *Reports on Progress in Physics*, 1952, 15, p. 275.

(189) KOMPFNER, R.: 'On the Operation of the Travelling-Wave Tube at Low Level', *Journal of the British Institution of Radio Engineers*, 1950, 10, p. 283.

(190) KOMPFNER, R., and WILLIAMS, N. T.: 'Backward-Wave Tubes', *Proceedings of the Institute of Radio Engineers*, 1953, 41, p. 1602.

(191) KREUCHEN, K. H., AULD, B. A., and DIXON, N. E.: 'A Study of the Broad-Band Frequency Response of the Multicavity Klystron Amplifier', *Journal of Electronics*, 1957, 2, p. 529.

(192) LABUS, J.: 'Space-Charge Waves along Magnetically Focused Electron Beams', *Proceedings of the Institute of Radio Engineers*, 1957, 45, p. 854.

(193) LABUS, J.: 'Optimum Amplification in the Travelling-Wave Valve with Helix', *Archiv der Elektrischen Übertragung*, 1952, 6, p. 1.

(194) LABUS, J.: 'Modulation Properties of the Reflex Klystron Generator', *Fernmeldetechnische Zeitschrift*, 1954, 7, p. 562.

(195) LACY, P. D.: 'Helix-Coupled Traveling-Wave Tubes', *Electronics*, November, 1954, p. 132.

(196) LAFFERTY, J. M.: 'A Millimetre Wave Reflex Oscillator', *Journal of Applied Physics*, 1946, 17, p. 1061.

(197) LAFFERTY, J. M.: 'Velocity-Modulated Reflex Oscillator', *Proceedings of the Institute of Radio Engineers*, 1947, 35, p. 913.

(198) LALLY, P. M.: 'A Space Harmonic Traveling-Wave Amplifier', *Proceedings of the National Electronics Conference* (*Chicago*), 1952, 8, p. 73.

(199) LAPIERRE, A. B.: 'Development of Travelling Wave Amplifiers with a Helical Coil: General Results', *Annales des Télécommunications*, 1948, 3, p. 259.

(200) LAPLUME, J.: 'Theory of the Travelling-Wave Valve', *L'Onde Électrique*, 1949, 29, p. 66.

(201) LAPOSTOLLE, P.: 'Study of the Different Waves that can be propagated along a Line of Interaction with an Electron Beam. Application to the Theory of the Travelling-Wave Amplifier', *Annales des Télécommunications*, 1948, 3, pp. 57 and 85.

(202) LAPOSTOLLE, P.: 'Interaction Phenomena in the Travelling-Wave Valve: Theory and Experimental Verification', *ibid.*, 1948, 3, p. 265.

(203) LAPUKHIN, V. M., and SAMORODOV, YU. D.: 'A Graphical Method for Investigating Travelling-Wave Valves', *Zhurnal Tekhnicheskoi Fiziki*, 1955, **25**, p. 1265.
(204) LAVOO, N. T.: 'Transadmittance and Input Conductance of a Lighthouse Triode at 3,000 Megacycles', *Proceedings of the Institute of Radio Engineers*, 1947, **35**, p. 1248.
(205) LAWSON, J. D., BARTON, R. S., GUBBINS, T. F., MILLAR, W., and ROGERS, P. S.: 'The Design and Performance of a High-Power Demountable Klystron Amplifier for X-Band', *Journal of Electronics*, 1955, **1**, p. 333.
(206) LEARNED, V., and VERONDA, C.: 'Recent Developments in High-Power Klystron Amplifiers', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 465.
(207) LINDER, E. G., and SPROULL, R. L.: 'The Maximum Efficiency of Reflex-Klystron Oscillators', *ibid.*, 1947, **35**, p. 241.
(208) LITTLE, J. B.: 'Amplification at 6-Millimetre Wavelength', *Bell Laboratories Record*, 1951, **29**, p. 14.
(209) LOB, C. G., and HOLHOUSER, D. F.: 'Radial-Beam Velocity-Modulated Microwave Tube', *Proceedings of the National Electronics Conference* (*Chicago*), 1949, **5**, p. 403.
(210) LOSHAKOV, L. N.: 'Approximate Calculation of the Propagation Constants of Transmission Lines in the Presence of an Electron Beam', *Zhurnal Tekhnicheskoi Fiziki*, 1956, **26**, p. 809.
(211) LOSHAKOV, L. N.: 'Theory of the Propagation of Waves in an Electron Beam', *ibid.*, 1952, **22**, p. 193.
(212) LOUISELL, W. H., and PIERCE, J. R.: 'Power Flow in Electron Beam Devices', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 425.
(213) MACFARLANE, G. G., and WOODWARD, A. M.: 'Small-Signal Theory of Wave Propagation in a Uniform Electron Beam', *Proceedings I.E.E.*, Paper No. 1002, September, 1950 (97, Part III, p. 322).
(214) MACFARLANE, G. G., and HAY, H. G.: 'Wave Propagation in a Slipping Stream of Electrons: Small Amplitude Theory', *Proceedings of the Physical Society*, 1950, **63**, p. 409.
(215) MCNALLY, J. O., and SHEPHERD, W. G.: 'Reflex Oscillators for Radar Systems', *Proceedings of the Institute of Radio Engineers*, 1947, **35**, p. 1424.
(216) MCWHIRT, R. B.: 'Wide-Band Amplifier for UHF Receivers', *Electronics*, December, 1955, p. 158.
(217) MATHEWS, A. R., SAH, C. T., and SPANGENBERG, K. R.: 'A Travelling Wave Tube with Interchangeable External Slow Wave Circuits', *L'Onde Électrique*, 1957, **37**, p. 65.
(218) MATRICON, M.: 'Development of a Demountable Klystron for the Generation of Millimetre Waves', *Revue Technique Compagnie Française Thomson-Houston*, 1951, **16**, p. 45.
(219) MATSUO, Y.: 'Multi-Beam Velocity-type Frequency Multiplier', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 101.
(220) MELZER, B.: 'Two-Cavity Klystron', *Wireless Engineer*, 1949, **26**, p. 365.
(221) MELZER, B.: 'Notes on the Multi-Reflection Klystron', *Electronic and Radio Engineer*, 1957, **34**, p. 109.
(222) MIHRAN, T. G.: 'The Duplex Travelling-Wave Klystron', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 308.
(223) MOHRAN, T. G.: 'Scalloped Beam Amplification', *Journal of Applied Physics*, 1954, **25**, p. 1341.
(224) MIHRAN, T. G.: 'Scalloped Beam Amplification', *Transactions of the Institute of Radio Engineers*, 1956, **ED–3**, No. 1, p. 32.
(225) MIKHALEVSKI, V. S., and VENEROVSKI, D. N.: 'Generation of Electromagnetic Oscillations by means of a Travelling-Wave Valve with an External Helix', *Zhurnal Tekhnicheskoi Fiziki*, 1955, **25**, p. 812.
(226) MIKHALEVSKI, V. S., and VENEROVSKI, D. N.: 'Generation of Electromagnetic Oscillations by a Travelling-Wave Valve with an External Sectionalised Helix', *ibid.*, 1956, **26**, p. 526.
(227) MIKHALEVSKI, V. S., DOLGANOV, A. G., and IVANOVA, V. D.: 'Generation of Electromagnetic Oscillations by means of a Travelling-Wave Valve with a Twin-Helix Coaxial Line', *Radiotekhnika i Elektronika*, 1956, **1**, p. 1383.
(228) MILLMAN, S.: 'A Spatial-Harmonic Traveling-Wave Amplifier for Six-Millimetre Wavelength', *Proceedings of the Institute of Radio Engineers*, 1951, **39**, p. 1035.
(229) MILLMAN, S.: 'Spatial-Harmonic Traveling-Wave Amplifier', *Bell Laboratories Record*, 1952, **30**, p. 413.
(230) MINAKOVA, I. I., and STEPANOVA, N. V.: 'Synchronisation of a Reflex Klystron', *Radiotekhnika i Elektronika*, 1956, **1**, p. 805.
(231) MIWA, T., KOYAMA, J., MISHIMA, M., and YANAOKA, J.: 'Helix Type Travelling-Wave Valve for 24 Gc/s', *Le Vide*, 1957, **12**, p. 49.
(232) MOLL, J. L., and WILMARTH, R. W.: 'A Wide Range Oscillator in the Range from 8 000 to 15 000 Mc/s', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 813.
(233) MORENO, T., and JEPSEN, R. L.: 'Hysteresis in Klystron Oscillators', *ibid.*, 1955, p. 344.
(234) MOTZ, H.: 'An Analysis of Klystron Reflector Performance', *Journal I.E.E.*, 1948, **95**, Part III, p. 295.
(235) MOURIER, G.: 'Contribution to the Large-Signal Theory of Travelling-Wave Tubes', *Annales de Radioélectricité*, 1956, **11**, p. 271.
(236) MULLEN, J. A.: 'A Power Series Solution of the Traveling-Wave Tube Equations', *Transactions of the Institute of Radio Engineers*, 1957, **ED–4**, No. 2, p. 159.
(237) MULLER, M.: 'Traveling Wave Amplifiers and Backward-Wave Oscillators', *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 1651.
(238) MULLER, R.: 'Space Harmonics in Electron Streams', *Archiv der Elektrischen Übertragung*, 1956, **10**, p. 585.
(239) MURAKAMI, T.: 'A Study of Grounded Grid UHF Amplifiers', *RCA Review*, 1951, **12**, p. 682.
(240) NAKAMURA, Y., MIWA, T., and HASEGAWA, Y.: 'Microwave Triode B–26 for 4 000 Mc/s Operation', *L'Onde Électrique*, 1957, **37**, p. 86.
(241) NALOS, E. J.: 'A Hybrid-Type Traveling-Wave Tube for High-Power Pulsed Amplification', *Transactions of the Institute of Radio Engineers*, 1958, **ED–5**, No. 3, p. 161.
(242) NERGAARD, L. S.: 'Analysis of a Simple Model of a Two-Beam Growing-Wave Tube', *RCA Review*, 1948, **9**, p. 585.
(243) NISHIO, H., NEMOTO, T., and MURAKAMI, H.: 'Some Problems on Disk-Sealed Planar Triodes', *Le Vide*, 1957, **12**, p. 7.
(244) NOLAND, J. A., and LEPIC, R. E.: 'Backward-Wave Oscillators for the 17 to 41 kMc/s Band', *Sylvania Technologist*, 1957, **10**, p. 13.
(245) NORDSIEK, A.: 'Theory of the Large-Signal Behaviour of Traveling-Wave Amplifiers', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 630.
(246) OHLSEN, L. H. VON: 'The Small Signal Performance of the 416B planar triode between 60 and 4000 Mc/s', *Transactions of the Institute of Radio Engineers*, 1954, **ED–1**, No. 4, p. 189.
(247) OLVING, S.: 'Amplification of the Travelling-Wave Tube at High Beam Current', *Chalmers Tekniska Högskolas Handlingar*, 1955, No. 157, p. 1.
(248) PALLUEL, P.: 'Recent Developments in O-Type Carcinotron Tubes', *Annales de Radioélectricité*, 1956, **11**, p. 145, and *L'Onde Électrique*, 1956, **36**, p. 318.
(249) PALLUEL, P., and GOLDBERGER, A. K.: 'The O-Type Carcinotron Tube', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 333.
(250) PAN, W. Y.: 'Investigation of UHF Television Amplifier Tubes', *RCA Revi* 1954, **15**, p. 27, and *Transactions of the Institute of Radio Engineers*, 19 ED–1, No. 1, p. 8.
(251) PARZEN, P., and GOLDSTEIN, L.: 'Effect of Hydrostatic Pressure in an Electr Beam on the Operation of Travelling-Wave Devices', *Journal of Appli Physics*, 1951, **22**, p. 398.
(252) PARZEN, P.: 'Theory of Space-Charge Waves in Cylindrical Waveguides w Many Beams', *Electrical Communication*, 1951, **28**, p. 217.
(253) PARZEN, P.: 'Space-Charge Wave Propagation in a Cylindrical Electron Bea of Finite Lateral Extension', *Journal of Applied Physics*, 1952, **23**, p. 215.
(254) PASCHKE, F.: 'Theory of Coupled Space-Charge Waves', *Frequenz*, 1955, p. 273.
(255) PASCHKE, F.: 'The Reciprocity of the Coupling in Travelling-Wave Valve *Archiv der Elektrischen Übertragung*, 1957, **11**, p. 137.
(256) PAUL, H.: 'The Frequency Dependence of Self-Excited Microwave Oscillat under Complex Load', *Elektronische Rundschau*, 1956, **10**, pp. 146 and 167.
(257) PEARCE, A. F., and MAYO, B. J.: 'The CV35—a Velocity-Modulation Reflecti Oscillator for Wavelengths of about 10 cm', *Journal I.E.E.*, 1946, **93**, Part III No. 5, p. 918.
(258) PEARCE, A. F.: 'A Velocity-Modulation Reflection Oscillator for Wavelengt of about 3·2 cm', *ibid.*, 1948, **95**, Part III, p. 415.
(259) PEARCE, A. F., KREUCHEN, K. H., BARON, C., HOULDING, N., and RATCLIFF, S 'Plug-In Reflex Klystrons for Microwaves', *Journal of Electronics and Contr* 1957, **3**, p. 535.
(260) PEEK, S. C.: 'A Disc-Seal Triode as a UHF Amplifier', *Sylvania Technologi* 1953, **6**, p. 81.
(261) PENNING, F. M.: 'Velocity-Modulation Valves', *Philips Technical Review*, 194 **8**, p. 214.
(262) PETER, R. W., BLOOM, S., and RUETZ, J. A.: 'Space-Charge-Wave Amplificati along an Electron Beam by Periodic Change of the Beam Impedance', *RC Review*, 1954, **15**, p. 113.
(263) PIDDINGTON, J. H.: 'Growing Electric Space-Charge Waves and Haef Electron-Wave Tube', *Physical Review*, 1956, **101**, p. 14.
(264) PIERCE, J. R., and SHEPHERD, W. G.: 'Reflex Oscillators', *Bell System Technic Journal*, 1947, **26**, p. 460.
(265) PIERCE, J. R., and FIELD, L. M.: 'Traveling-Wave Tubes', *Proceedings of t Institute of Radio Engineers*, 1947, **35**, p. 108.
(266) PIERCE, J. R.: 'Theory of the Beam-type Traveling-Wave Tube', *ibid.*, 194 **35**, p. 111.
(267) PIERCE, J. R.: 'Transverse Fields in Traveling-Wave Tubes', *Bell Syste Technical Journal*, 1948, **27**, p. 732.
(268) PIERCE, J. R.: 'Effect of Passive Modes in Traveling-Wave Tubes', *Proceedin of the Institute of Radio Engineers*, 1948, **36**, p. 993.
(269) PIERCE, J. R., and HEBENSTREIT, W. B.: 'A New Type of High Frequen Amplifier', *Bell System Technical Journal*, 1949, **28**, p. 33.
(270) PIERCE, J. R., and WAX, N.: 'A Note on Filter-type Traveling-Wave Amplifier *Proceedings of the Institute of Radio Engineers*, 1949, **37**, p. 622.
(271) PIERCE, J. R.: 'Traveling-Wave Tubes: Parts 1 to 4', *Bell System Technic Journal*, 1950, **29**, pp. 1, 189, 390 and 608.
(272) PIERCE, J. R.: 'Travelling-Wave Tubes' (Van Nostrand, New York, 1950).
(273) PIERCE, J. R.: 'Millimetre Waves', *Physics Today*, November, 1950, **3**, p. 24.
(274) PIERCE, J. R.: 'Waves in Electron Streams and Circuits', *Bell System Techni Journal*, 1951, **30**, p. 626.
(275) PIERCE, J. R.: 'Some Recent Advances in Microwave Tubes', *Proceedings the Institute of Radio Engineers*, 1954, **42**, p. 1735.
(276) PIERCE, J. R.: 'The Wave Picture of Microwave Tubes', *Bell System Techni Journal*, 1954, **33**, p. 1343.
(277) PIERCE, J. R.: 'Interaction of Moving Charges with Wave Circuits', *Journal Applied Physics*, 1955, **26**, p. 627.
(278) PIERCE, J. R.: 'Double Stream Amplifiers', *Proceedings of the Institute of Ra Engineers*, 1949, **37**, p. 980.
(279) PIERCE, J. R.: 'Coupling of Modes of Propagation', *Journal of Applied Physi* 1954, **25**, p. 179.
(280) PIERCE, J. R., and WALKER, L. R.: 'Growing Electric Space-Charge Wave *Physical Review*, 1956, **104**, p. 306.
(281) PIERCE, J. R., and WALKER, L. R.: 'Growing Waves due to Transverse Velocitie *Bell System Technical Journal*, 1956, **35**, p. 109.
(282) POSCHL, K.: 'Influencing Space-Charge Waves of Fluctuating Beams Resonator Circuits', *Frequenz*, 1954, **8**, p. 284.
(283) PRADHAM, T.: 'Plasma Oscillations in a Steady Magnetic Field: Circula Polarised Electromagnetic Modes', *Physical Review*, 1957, **107**, p. 1222.
(284) PREISACH, F., and ZAKARIAS, I.: 'Input Conductance', *Wireless Engineer*, 194 **17**, p. 147.
(285) PREIST, D. H.: 'Annular Circuits for High Power Multiple Tube Radial Fr quency Generators at VHF and UHF', *Proceedings of the Institute of Ra Engineers*, 1950, **38**, p. 515.
(286) PRIEST, D. H., MURDOCH, C. E., and WOERNER, J. J.: 'High Power Klystro at UHF', *ibid.*, 1953, **41**, p. 20.
(287) RAMO, S.: 'Space Charge and Field Waves in an Electron Beam', *Physi Review*, 1939, **56**, p. 276.
(288) REED, E. D.: 'A Coupled Resonator Reflex Klystron', *Bell System Techni Journal*, 1953, **32**, p. 715.
(289) REED, E. D.: 'A Tunable, Low-Voltage Reflex Klystron for Operation in t 50–60 kMc/s Band', *ibid.*, 1955, **34**, p. 563.
(290) RIGROD, W. W., and LEWIS, J. A.: 'Wave Propagation along a Magnetical Focussed Cylindrical Electron Beam', *ibid.*, 1954, **33**, p. 399.
(291) ROBERTSON, S. D.: 'Electronic Admittances of Parallel-Plane Electron Tu at 4,000 Mc/s', *ibid.*, 1949, **28**, p. 619.
(292) ROBERTSON, S. D.: 'An Experimental Broad-Band Helix Traveling-Wa Amplifier for Millimetre Wavelengths', *Transactions of the Institute of Ra Engineers*, 1954, **MTT–2**, No. 3, p. 48.
(293) ROBINSON, F. N. H.: 'Travelling-Wave Tubes with Dispersive Helices', *Wirel Engineer*, 1951, **28**, p. 110.
(294) ROBINSON, F. N. H.: 'Travelling Wave Tubes', *Research, London*, 1956, **9**, p.
(295) RODENHUIS, K.: 'Construction and Circuit of a 4 000 Mc/s Triode with Cathode', *Le Vide*, 1957, **12**, p. 23.
(296) ROGERS, D. C.: 'Traveling Wave Amplifier for 6 to 8 Centimetres', *Electri Communication*, 1949, **26**, p. 144.
(297) ROWE, J. E.: 'Design Information on Large-Signal Traveling-Wave Amplifie *Proceedings of the Institute of Radio Engineers*, 1956, **44**, pp. 200 and 818.
(298) ROWE, J. E.: 'A Large-Signal Analysis of the Traveling-Wave Amplifi Theory and General Results', *Transactions of the Institute of Radio Enginee* 1956, **ED–3**, No. 1, p. 39.
(299) RYDBECK, O. E. H.: 'The Theory of the Travelling-Wave Tube', *Erics Technics*, 1948, No. 46, p. 3.
(300) RYDBECK, O. E. H.: 'On the Excitation of Different Space-Charge Wave Modes Travelling-Wave Tubes', *Archiv der Elektrischen Übertragung*, 1953, **7**, p. 4
(301) RYDBECK, O. E. H., and AGDUR, B.: 'The Propagation of Electron Space-Cha Waves in Waveguides and Tubes with Periodic Structure', *L'Onde Électriq* 1954, **34**, p. 499, and *Chalmers Tekniska Högskolas Handlingar*, 1954, **138**, p

(12) RYNN, N.: 'Analysis of Coupled-Structure Traveling Wave Tubes', *Transactions of the Institute of Radio Engineers*, 1957, ED–4, No. 2, p. 172.

(13) SATO, K., KOBAYASHI, D., KONDO, A., and KOYAMA, J.: 'Development of Travelling-Wave Tubes for 4 000 Mc/s Band', *Reports of the Electrical Communication Laboratory (Japan)*, 1955, 3, p. 11.

(14) SCHNITZER, H.: 'Amplification and Generation of Microwave Oscillations with Travelling-Wave Valves', *Funk und Ton*, 1951, 5, p. 143.

(15) SCHNITZER, H.: 'Production of Direction-Dependent Electronic Attenuator in Travelling-Wave Valves', *Fernmeldetechnische Zeitschrift*, 1951, 4, p. 301.

(16) SCHNITZER, H., and WEBER, D.: 'Experimental Verification of Small-Signal Theory for the Travelling-Wave Valve with Helix', *Archiv der Elektrischen Übertragung*, 1952, 6, p. 369.

(17) SCHNITZER, H., and WEBER, D.: 'Measurements on a 10-W Helix-Type Travelling-Wave Valve for 15 cm Wavelength', *Fernmeldetechnische Zeitschrift*, 1953, 6, p. 66.

(18) SCHNITZER, H.: 'Measurement of the Cross-Sectional Non-Uniformity of the Electronic Beam in a Helix-Type Travelling-Wave Valve', *Archiv der Elektrischen Übertragung*, 1953, 7, p. 415.

(19) SCHUMANN, W. O.: 'Waves and Plasma', *Elektron Wissen Technik*, 1951, 5, p. 279.

(10) SCOTTO, M., and PARZEN, P.: 'The Electronic Theory of Tape-Helix Traveling-Wave Structures', *Transactions of the Institute of Radio Engineers*, 1955, ED–2, No. 4, p. 19.

(11) SEVERIN, H.: 'Generation of Oscillators at Centimetre Wavelengths', *Technische Mitteilungen PTT*, 1951, 29, p. 466.

(12) SHEPHERD, W. G.: 'Efficiency of Reflex Oscillators', *Proceedings of the National Electronics Conference (Chicago)*, 1949, 5, p. 500.

(13) SHEVCHIK, V. N.: 'Analysis of the Energy Interchange between an Electron Stream and an Electromagnetic Wave', *Radiotekhnika i Elektronika*, 1957, 2, p. 104.

(14) SHEVCHIK, W. N., and ZHARKOV, Y. D.: 'The Cascade Bunching of Electrons Applied to the Analysis of the Interaction of Electron Stream and Travelling Magnetic-Wave', *ibid.*, 1957, 2, p. 237.

(15) SHEVCHIK, W. N., SUSLOV, S. A., and ZHARKOV, Y. D.: 'Investigation of a Special Type of Reflex Klystron', *Zhurnal Tekhnicheskoi Fiziki*, 1957, 27, p. 377.

(16) SHULMAN, C., and HEAGY, M. S.: 'Small-Signal Analysis of Traveling-Wave Tube', *RCA Review*, 1947, 8, p. 585.

(17) SIEGMAN, A. E., and JOHNSON, H. R.: 'Suppression of Backward-Wave Oscillation by Filter Helix Methods', *Transactions of the Institute of Radio Engineers*, 1955, ED–2, No. 2, p. 20.

(18) SIEKANOWICZ, W. W., and STERZER, F.: 'A Developmental Wide Band S-band Traveling-Wave Amplifier utilizing Periodic Permanent Magnets', *Proceedings of the Institute of Radio Engineers*, 1956, 44, p. 55.

(19) SIGRIST, W.: 'Observations on Some Electronic Fundamentals of Microwave Valves', *Bulletin Schweizer Electrotechnik*, 1950, 41, p. 35.

(20) SPEAKS, F. A.: 'Power Amplifier Klystron for UHF Transmission', *Transactions of the Institute of Radio Engineers*, 1956, CS–4, p. 69.

(21) STAINSBY, A. G.: 'Backward Wave Oscillators', *Electronic Engineering*, 1958, 30, p. 329.

(22) STEPHENSON, J. G.: 'Designing Stable Triode Microwave Oscillators', *Electronics*, March, 1955, p. 184.

(23) SULLIVAN, J. W.: 'A Wide-Band Voltage-Tunable Oscillator', *Proceedings of the Institute of Radio Engineers*, 1954, 42, p. 1658.

(24) SWIFT, J.: 'Disc-Seal Circuit Techniques', *Proceedings of the Institution of Radio Engineers, Australia*, 1955, 16, pp. 205, 247 and 295.

(25) SWIFT, J.: 'Disc-Seal Oscillators—Parts I and II', *Journal of the British Institution of Radio Engineers*, 1955, 15, p. 607, and 1956, 16, p. 95.

(26) SZULKIN, P.: 'The Excitation of a Cavity Resonator by a Density-Modulated Electron Beam Passing through the Entire Resonator Cross-Section', *Archiwum Elektrotechniki*, 1956, 5, p. 149.

(27) TIEN, P. K., and FIELD, L. M.: 'Space-Charge Waves in an Accelerated Electron Stream for Amplification of Microwave Signals', *Proceedings of the Institute of Radio Engineers*, 1952, 40, p. 688.

(28) TIEN, P. K., WALKER, L. R., and WOLONTIS, V. M.: 'A Large-Signal Theory of Traveling-Wave Amplifiers', *ibid.*, 1955, 43, p. 260.

(29) TIEN, P. K.: 'A Large-Signal Theory of Traveling-Wave Amplifiers', *Bell System Technical Journal*, 1956, 35, p. 349.

(30) THOMSON, J.: 'Microwave Valves', *Journal I.E.E.*, 1956, 2, p. 661.

(31) TONKS, L., and LANGMUIR, I.: 'Oscillations in Ionised Gases', *Physical Review*, 1929, 33, p. 195.

(32) TORSTENSSON, L.: 'Reflex Klystron for Grid Pulse Operation within the Frequency Range 2 000–5 000 Mc/s', *Ericsson Technics*, 1954, 10, p. 297.

(33) TREMLETT, C. A., and WILLIAMS, A. D.: 'Triodes and Tetrodes for UHF-SHF Operation', *Electronic Engineering*, 1958, 30, p. 335.

(34) UDA, S., and IKEUCHI, J.: 'Equivalent Networks of Klystrons', *Technological Reports (Tohoku University)*, 1950, 14, p. 117.

(35) UDAGAWA, K.: 'Theoretical Study of the Traveling-Wave Tube', *Reports of the Electrical Communication Laboratory (Japan)*, 1954, 2, p. 34.

(36) VAINSHTEIN, L. A.: 'Electron Waves in Retarding Systems', *Zhurnal Tekhnicheskoi Fiziki*, 1956, 26, p. 126.

(37) VALLEY, G. E., and WALLMAN, H.: 'Vacuum Tube Amplifiers' (McGraw-Hill, New York, 1948).

(38) VARIAN, R. H.: 'Recent Developments in Klystrons', *Electronics*, April, 1952, p. 112.

(39) VASILLEV, E. I., and LOPUKHIN, V. M.: 'Self-Excitation of a Decelerating System', *Zhurnal Tekhnicheskoi Fiziki*, 1951, 21, p. 527.

(40) VASILLEV, E. I., and LOPUKHIN, V. M.: 'On the Theory of the Helix-Type Travelling-Wave Valve', *ibid.*, 1952, 22, p. 1838.

(41) VAUGHAN, J. R. M.: 'Klystron Modulation and Schlomilch Series', *Journal of Electronics*, 1956, 1, p. 430.

(42) VEITH, W.: 'The Carcinotron, an Electrically Tuned Microwave Generator', *Fernmeldetechnische Zeitschrift*, 1954, 7, p. 554.

(43) VINCENT, G., YEOU, T., and LABARDERIE, J.: 'Millimetre Range Carcinotrons', *L'Onde Électrique*, 1957, 37, p. 65.

(44) WADE, G., and RYNN, N.: 'Coupled Helices for use in Traveling Wave Tubes', *Transactions of the Institute of Radio Engineers*, 1955, ED–2, p. 15.

(45) WALKER, L. R.: 'Power Flow in Electron Beams', *Journal of Applied Physics*, 1955, 26, p. 1031.

(46) WALKER, L. R.: 'Stored Energy and Power Flow in Electron Beams', *ibid.*, 1954, 25, p. 615.

(47) WALKER, L. R.: 'Starting Currents in the Backward-Wave Oscillator', *ibid.*, 1953, 24, p. 854.

(48) WALLAUSCHEK, R.: 'Experimental Determination of the Characteristics of Travelling-Wave Amplifiers: Results obtained', *Annales des Télécommunications*, 1948, 3, p. 300.

(49) WANG, C. C., PIERCE, J. R., and TIEN, P. K.: 'Power Flow and Equivalent Circuits of Traveling-Wave Tubes', *Proceedings of the Institute of Radio Engineers*, 1954, 42, p. 1701.

(50) WANG, C. C.: 'Linear Beam Tube Theory', *Transactions of the Institute of Radio Engineers*, 1957, ED–4, No. 1, p. 92.

(351) WARNECKE, R., KLEEN, W., DOEHLER, O., and HUBER, H.: 'Electron Bunching in a Velocity-Modulation Valve by means of a Travelling-Wave Device', *Comptes Rendus (Paris)*, 1949, 229, p. 648.

(352) WARNECKE, R., DOEHLER, O., and KLEEN, W.: 'Electron Beams and Electromagnetic Waves', *Wireless Engineer*, 1951, 28, p. 167.

(353) WARNECKE, R., and GUENARD, P.: 'Velocity Modulation Valves' (Gauthier Villars, Paris, 1951).

(354) WARNECKE, R.: 'Some Recently Obtained Results on Valves for Microwave Frequencies', *Annales de Radioélectricité*, 1954, 9, p. 107.

(355) WARNECKE, R.: 'The Principal Results Achieved in the Field of Microwave Valves', *Le Vide*, 1956, 11, p. 217.

(356) WARNECKE, R., GUENARD, P., and DOEHLER, O.: 'Fundamental Phenomena in Travelling Wave Tubes', *L'Onde Électrique*, 1954, 34, p. 323.

(357) WATERS, W. E.: 'Space-Charge Effects in Klystrons', *Transactions of the Institute of Radio Engineers*, 1957, ED–4, No. 1, p. 49.

(358) WATKINS, D. A.: 'Effects of Velocity Distribution in a Modulated Electron Beam', *Journal of Applied Physics*, 1952, 23, p. 568.

(359) WATKINS, D. A., and RYNN, N.: 'Effect of Velocity Distribution on Travelling-Wave Tube Gain', *ibid.*, 1954, 25, p. 1375.

(360) WEBBER, S. E.: '1,000 Watt Traveling-Wave Tube', *Electronics*, June, 1950, p. 100.

(361) WEBBER, S. E.: 'Ballistic Analysis of a Two-Cavity Finite Beam Klystron', *Transactions of the Institute of Radio Engineers*, 1958, ED–5, No. 2, p. 98.

(362) WEBBER, S. E.: 'Electron Bunching and Energy Exchange in a Traveling-Wave Tube', *ibid.*, 1957, ED–4, p. 87.

(363) WEBER, D.: 'A Computational Contribution to the Linear Theory of the Travelling-Wave Valve with Low Gain', *Archiv der Elektrischen Übertragung*, 1954, 8, p. 341.

(364) WEBSTER, D. L.: 'The Theory of Klystron Generators', *Journal of Applied Physics*, 1939, 10, p. 864.

(365) WEGLEIN, R. D.: 'Backward-Wave Oscillator Starting Conditions', *Transactions of the Institute of Radio Engineers*, 1957, ED–4, No. 2, p. 177.

(366) WEHNER, G.: 'Electron Plasma Oscillations', *Journal of Applied Physics*, 1951, 22, p. 761.

(367) WEINSTEIN, M., and FOERSTER, H. M. VAN: 'Space-Charge Effects in Dense, Velocity-Modulated Electron Beams', *ibid.*, 1956, 27, p. 344.

(368) WHINNERY, J. R., WATSON, W. H., and JORY, H. R.: 'Cascade Backward-Wave Amplifiers', *L'Onde Électrique*, 1957, 37, p. 86.

(369) WHINNERY, J. R.: 'Traveling-Wave-Tube Oscillators', *Transactions of the Institute of Radio Engineers*, 1953, ED–2, p. 11.

(370) WILHELMSSEN, H.: 'The Interaction between an Obliquely Incident Plane Electromagnetic Wave and an Electron Beam: Part I', *Chalmers Tekniska Högskolas Handlingar*, 1954, No. 155, p. 1.

(371) WILLENBROCK, F. K., and COOKE, S. P.: 'Interaction of a Spiral Electron Beam and a Resonant Microwave Cavity', *Journal of Applied Physics*, 1950, 21, p. 114.

(372) WILMARTH, R. W., and BOYCHENKO, O. G.: 'High Power S-Band Pulsed Travelling Wave Tube', *L'Onde Électrique*, 1957, 37, p. 86.

(373) WILMARTH, R. W., and MOLL, J. L.: 'A Wide-Range Oscillator in the Range from 8,000 to 15,000 Mc/s', *Proceedings of the Institute of Radio Engineers*, 1952, 40, p. 813.

(374) WOOTTON, D. J., and PEARCE, A. F.: 'A Reflex Klystron for the 8–9 mm Band', *Proceedings I.E.E.*, Monograph No. 143 R, August, 1955 (103 C, p. 104).

(375) WORSLEY, P. K.: 'Electron Tubes for Microwave Applications—a Survey of Available Types', *British Communications and Electronics*, 1956, 3, pp. 606 and 668.

(376) YADAVALLI, S. V.: 'Application of the Potential Analogue in Multicavity Klystron Design and Operation', *Proceedings of the Institute of Radio Engineers*, 1957, 45, p. 1286.

(377) ZITO, G.: 'Reflex Klystron Performance', *Zeitschrift für angewandte Mathematik und Physik*, 1954, 5, p. 252.

(378) 'Microminiature Tube for UHF and High Temperature', *Electrical Manufacturing*, 1955, 56, p. 154.

(379) 'I.R.E. Standards on Electron Devices: Definitions of Terms related to Microwave Tubes (Klystrons, Magnetrons and Travelling Wave Tubes) 1956', *Proceedings of the Institute of Radio Engineers*, 1956, 44, p. 346.

Part 2

(380) ABELE, M.: 'The Čerenkov Effect in Optics and in the Domain of Microwaves', *Il Nuovo Cimento*, 1952, 9, p. 207.

(381) ADLER, R., KROMHOUT, O. M., and CLAVIER, P. A.: 'Transverse-Field Traveling-Wave Tubes with Periodic Electrostatic Focusing', *Proceedings of the Institute of Radio Engineers*, 1956, 44, p. 82.

(382) AIKIN, A. W.: 'A Proposed Method of Investigating the Field Distribution in a Magnetron', *Journal of Scientific Instruments*, 1955, 32, p. 152.

(383) AKHIEZER, A. I., LYNBARSKI, G. Y., and FAINBERG, Y. B.: 'Radiation from a Charged Particle moving through Coupled Resonators', *Zhurnal Tekhnicheskoi Fiziki*, 1955, 25, p. 2526.

(384) AKHIEZER, A. I., and FAINBERG, Y. B.: 'Interaction of Coupled Electromagnetic Resonators with a Beam of Charged Particles', *ibid.*, 1955, 25, p. 2516.

(385) ASKAR' YAN, G. A.: 'Pulsed Coherent Generation of Millimetre Waves by Nonrelativistic Electron Bunches', *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki*, 1956, 30, p. 584.

(386) ASKAR' YAN, G. A.: 'The Radiation of an Accelerated Moving Electrical Image of a Uniformly Moving Charge', *ibid.*, 1955, 29, p. 3.

(387) ASKAR' YAN, G. A.: 'On the Generation of Radio Waves in the Millimetre Band by the Passage of a Bunch of Electrons through a Retarding Medium', *ibid.*, 1954, 27, p. 6.

(388) ASHKIN, A.: 'A High-Power Rising-Sun Magnetron', *Physical Review*, 1946, 69, p. 701.

(389) BAKHRAKH, L. E.: 'Theory of the Multi-Segment Magnetron', *Zhurnal Tekhnicheskoi Fiziki*, 1952, 22, p. 1008.

(390) BECKER, G. E.: 'Dependence of Magnetron Operation on the Radial Centering of the Cathode', *Transactions of the Institute of Radio Engineers*, 1957, ED–4, No. 2, p. 126.

(391) BENNETOT, M. DE: 'Study of the Oscillation Modes of the M-Type Carcinotron: Parts 1 and 2', *Annales de Radioélectricité*, 1955, 10, p. 328, and 1956, 11, p. 230.

(392) BERNSTEIN, M. J., and KROLL, N. M.: 'Magnetron Research at Columbia Radiation Laboratory', *Transactions of the Institute of Radio Engineers*, 1954, MTT–2, No. 3, p. 33.

(393) BOGDANKEVICH, L. S., and BOLOTOVSKY, B. M.: 'The Movement of a Charge Parallel to the Axis of a Cylindrical Channel in a Dielectric', *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki*, 1957, 32, p. 6.

(394) BOOT, H. A. H., and HARVIE, R. B. R. S.: 'Charged Particles in a Non-Uniform Radio-Frequency Field', *Nature*, 1957, 180, p. 1187.

(395) BOOT, H. A. H., and RANDALL, J. T.: 'The Cavity Magnetron', *Journal I.E.E.*, 1946, 93, Part IIIA, p. 928.

(396) BORGNIS, F.: 'A Simple Approximate Formula for the Resonance Frequency of a Cavity Magnetron', *Zeitschrift für angewandte Physik*, 1950, 2, p. 278.

(397) BOYD, J. A.: 'The Mitron—an Interdigital Voltage-Tunable Magnetron', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 332.
(398) BREWER, G. R.: 'On the Focusing of High Current Electron Beams', *Journal of Applied Physics*, 1954, **25**, p. 243.
(399) BRILLOUIN, L.: 'A Theorem of Larmor and its Importance for Electrons in Magnetic-Fields', *Physical Review*, 1945, **67**, p. 260.
(400) BRISTOL, T. R., and GRIFFIN, G. J.: 'Voltage-Tuned Magnetron for F.M. Applications', *Electronics*, May, 1957, p. 162.
(401) BROSSART, J., and DOEHLER, O.: 'On the Properties of Valves using a Constant Magnetic Field. Part 3', *Annales de Radioélectricité*, 1948, **3**, p. 328.
(402) BROWN, W. C.: 'Description and Operating Characteristics of the Platinotron—a New Microwave Tube Device', *Proceedings of the Institute of Radio Engineers*, 1957, **47**, p. 1209.
(403) BUCK, D. C.: 'Stability of a Cylindrical Electron Beam in Nonsinusoidal Periodic Magnetic-Focusing Fields', *Transactions of the Institute of Radio Engineers*, 1957, **ED–4**, No. 1, p. 44.
(404) BUNEMAN, O.: 'Comments on Magnetron Theory, with Particular Reference to some Recent Publications', *Journal of Electronics*, 1955, **1**, p. 314.
(405) BUNEMAN, O.: 'Generation and Amplification of Waves in Dense Charged Beams under Crossed Fields', *Nature*, 1950, **165**, p. 474.
(406) BUNEMAN, O.: 'A Toroidal Magnetron', *Proceedings of the Physical Society*, 1950, **63**, p. 278.
(407) ČERENKOV, P.: 'Emission of Light from Pure Liquids under the Influence of Fast β-Rays', *Comptes Rendus de l'Académie des Sciences de l'URSS*, 1937, **14**, p. 101.
(408) CHANG, K. K. N.: 'Periodic-Magnetic-Field Focusing for Low-Noise Traveling-Wave Tubes', *RCA Review*, 1955, **16**, p. 423.
(409) CHANG, K. K. N.: 'Optimum Design of Periodic Magnet Structures for Electron Beam Focusing', *ibid.*, 1955, **16**, p. 65.
(410) CHANG, K. K. N.: 'Beam Focusing by Periodic and Complementary Fields', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 62.
(411) CHANG, K. K. N.: 'Biperiodic Electrostatic Focusing for High Density Electron Beams', *ibid.*, 1957, **45**, p. 1522.
(412) CHEN, T. S., and KOVACH, L.: 'Electron Beam Focusing in Three Anode Guns for Traveling-Wave Tubes', *Journal of Electronics and Control*, 1957, **3**, p. 287.
(413) CLOGSTON, A. M., and HEFFNER, H.: 'Focusing of an Electron Beam by Periodic Fields', *Journal of Applied Physics*, 1954, **25**, p. 436.
(414) COLEMAN, P. D., and SIRKIS, M. D.: 'The Harmodotron—A Beam Harmonic, High Order Mode Device for Producing Millimetre and Sub-millimetre Waves', *ibid.*, 1955, **26**, p. 1385.
(415) COLLINS, G. B.: 'Microwave Magnetrons' (McGraw-Hill, New York, 1947).
(416) COMBE, R., and FEIX, M.: 'Frequencies and Power of Waves Radiated by a Magnetic Undulator', *Comptes Rendus Hebdomadaires des Séances de Académie des Sciences* (*Paris*), 1953, **237**, p. 1660.
(417) COMBE, R., and FRELOT, T.: 'Production of Millimetre Waves by a Magnetic Undulator', *ibid.*, 1955, **241**, p. 1559.
(418) COOK, J. S., KOMPFNER, R., and YOCOM, W. H.: 'Slalom Focusing', *Proceedings of the Institute of Radio Engineers*, 1957, **45**, p. 1517.
(419) COOK, J. S., LOUISELL, W. H., and YOCOM, W. H.: 'Stability of an Electron Beam on a Slalom Orbit', *Journal of Applied Physics*, 1958, **29**, p. 583.
(420) COOK, J. S., KOMPFNER, R., and YOCOM, W. H.: 'Slalom Flow—A New Method of Periodic Electrostatic Focusing of Electron Beams', *L'Onde Électrique*, 1957, **37**, p. 65.
(421) COOLEY, J. P., and ROHRBAUGH, J. H.: 'The Production of Extremely Short Electromagnetic Waves', *Physical Review*, 1954, **67**, p. 296.
(422) COPPOLA, P. P., and HUGHES, R. C.: 'A New Pressed Dispenser Cathode', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 351.
(423) COSTE, J., and DELCROIX, J. L.: 'Plasma Oscillations and Resonance Frequencies in a Magnetron in the Brillouin State', *Comptes Rendus Hebdomadaires des Séance de l'Académie des Sciences* (*Paris*), 1956, **242**, p. 87.
(424) CRAWFORD, F. H., and HARE, M. D.: 'A Tuneable Squirrel Cage Magnetron', *Proceedings of the Institute of Radio Engineers*, 1947, **35**, p. 361.
(425) DANOS, M., GESCHWIND, S., LASHINSKY, H., and VAN TRIER, A.: 'Čerenkov Effect at Microwave Frequencies', *Physical Review*, 1953, **92**, p. 828.
(426) DANOS, M.: 'Čerenkov Radiation from Extended Electron Beams', *Journal of Applied Physics*, 1955, **26**, p. 2.
(427) DANOS, M., and LASHINSKY, H.: 'Millimetre Wave Generation by Čerenkov Radiation', *Transactions of the Institute of Radio Engineers*, 1954, **MTT–2**, No. 3, p. 21.
(428) DAVID, E. E.: 'R.F. Phase Control in Pulsed Magnetrons', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 669.
(429) DOEHLER, O.: 'On the Properties of Valves using a Constant Magnetic-Field: Parts 1 and 2', *Annales de Radioélectricité*, 1948, **3**, pp. 29 and 169.
(430) DOEHLER, O., BROSSART, J., and MOURIER, G.: 'On the Properties of Valves using a Constant Magnetic Field: Travelling-Wave Valves with Magnetic Field. Part 4', *ibid.*, 1950, **5**, p. 293.
(431) DOEHLER, O.: 'The Magnetron as a Travelling-Wave Valve', *Funk und Ton*, 1951, **5**, pp. 146 and 257.
(432) DOEHLER, O., and GUILBAUD, G.: 'Amplification of Electromagnetic Waves due to Displacement of an Electron Beam in Crossed Electric and Magnetic Fields bounded by Resistive Walls', *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* (*Paris*), 1954, **238**, p. 1784.
(433) EDER, F. X.: 'Čerenkov Radiation', *Funk und Ton*, 1949, **3**, p. 67.
(434) ELLIOTT, R. S.: 'Some Limitations on the Maximum Frequency of Coherent Oscillations', *Journal of Applied Physics*, 1952, **23**, p. 812.
(435) EPSZTEIN, B.: 'Influence of Space Charge on the Excitation Current of a Carcinotron-Type Magnetron Oscillator', *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* (*Paris*), 1955, **240**, p. 408.
(436) EPSZTEIN, B.: 'Effects of Space Charge in Crossed Field Valves', *ibid.*, 1957 **244**, p. 2902.
(437) ESPERSIN, G. A.: 'Low-Voltage Tuneable X-Band Magnetron Developments', *Tele-Tech*, 1951, **10**, pp. 30, 50, 70 and 84.
(438) FAINBERG, Y. B., and KHIZHNYAK, N. A.: 'Energy Loss by a Charged Particle Passing through a Laminar Dielectric: Part 1', *Zhurnal Eksperimentalnoi i teoreticheskoi Fiziki*, 1957, **32**, p. 883.
(439) FARAGO, P. S., and MARX, G.: 'Quantum Effects in the Interation between Free Electrons and Electromagnetic Fields', *Physical Review*, 1955, **99**, p. 1063, and *Acta Physica Hungarica*, 1954, **4**, p. 23.
(440) FARRANDS, J. L.: 'The Generation of Millimetre Waves', *Proceedings I.E.E.*, Monograph No. 112 R, December, 1954 (**102 C**, p. 98).
(441) FECHNER, P.: 'Study of the Magnetron in the Cut-off State', *Annales de Radioélectricité*, 1952, **7**, pp. 83 and 199.
(442) FECHNER, P.: 'Measurements of the Space Charge Resonance Frequencies in a Multi-Cavity Magnetron', *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* (*Paris*), 1950, **231**, p. 270.
(443) FECHNER, P.: 'Measurement of the Resonance Frequencies of the Space Charge in a Coaxial Magnetron', *ibid.*, 1950, **231**, p. 124.
(444) FEINSTEIN, J., and KINO, G. S.: 'The Large-Signal Behaviour of Crossed-Field Traveling-Wave Devices', *Proceedings of the Institute of Radio Engineers*, 1957, **45**, p. 1364.
(445) FEINSTEIN, J., and COLLIER, R. J.: 'A Magnetron Controlled by a Symmetrical Coupled TE_{011} Mode Cavity', *L'Onde Électrique*, 1957, **37**, p. 65.
(446) FISK, J. B., HAGSTRUM, H. D., and HARTMAN, P. L.: 'The Magnetron as Generator of Centimetre Waves: Parts 1 and 2', *Bell System Technical Journ* 1946, **25**, pp. 167 and 264.
(447) FRAENKEL, Z.: 'The Development of a Tunable C.W. Magnetron in the K-B Region', *Transactions of the Institute of Radio Engineers*, 1957, **ED–4**, No p. 271.
(448) FRANK, I., and TAMM, I.: 'Coherent Visible Radiation from Fast Electr Passing through Matter', *Comptes Rendus de l'Académie des Sciences l'URSS*, 1937, **14**, p. 109.
(449) FRITZ, K.: 'The Potentials in the Magnetron: Consequences in Desig *Fernmeldetechnische Zeitschrift*, 1954, **7**, p. 528.
(450) FRITZ, K.: 'Potentials and Electron Paths in Multi-Segment Magnetron *Archiv der Elektrischen Übertragung*, 1952, **6**, p. 211.
(451) FRITZ, K.: 'The State of Magnetron Development in Germany', *Funk und T* 1953, **7**, p. 133.
(452) GABOR, D., and SIMS, G. D.: 'Theory of the Pre-Oscillating Magnetron: Part *Journal of Electronics*, 1955, **1**, p. 25.
(453) GABOR, D.: 'Stationary Electron Swarms in Electromagnetic Fields', *Proceedi of the Royal Society*, 1944, **183A**, p. 436.
(454) GABOR, D., SIMS, G. D., and TWISS, R. Q.: 'Magnetron Theory', *Journal Electronics*, 1956, **1**, pp. 449 and 454.
(455) GERSHTEIN, G. M., and VITELS, G. L.: 'Widening of Oscillation Zones Decimetre-Waveband Magnetrons', *Radiotekhnika i Elektronika*, 19 **2**, p. 120.
(456) GINSBERG, V. L.: 'On the Radiation of an Electron Moving near a Dielectri *Comptes Rendus de l'Académie des Sciences de l'URSS*, 1947, **56**, p. 2.
(457) GINSBERG, V. L.: 'On the Utilisation of the Čerenkov Effect for the Radiati of Microwaves', *ibid.*, 1947, **56**, p. 2.
(458) GINSBERG, V. L., and FAIN, V. M.: 'On Quantum Effects occurring on Interacti of Electrons with High Frequency Fields in Resonant Cavities', *Zhurn Ekaperimentakoi i Teoreticheskoi Fisiki*, 1957, **32**, p. 162, and *Journal Experimental and Theoretical Physics* (*U.S.S.R.*), 1957 **5**, p. 123.
(459) GLAGOLEWA-ARKADIEWA, A.: 'Short Electromagnetic Waves of Waveleng up to 82 Microns', *Nature*, 1924, **113**, p. 640.
(460) GLASS, M. S.: 'Straight-Field Permanent Magnets of Minimum Weight f TWT Focusing', *Proceedings of the Institute of Radio Engineers*, 1957, 4 p. 1100.
(461) GLASS, R. C., SIMS, G. D., and STAINSBY, A. G.: 'Noise in Cut-off Magnetron *Proceedings I.E.E.*, Paper No. 1760 R, January, 1955 (**102 B**, p. 81).
(462) GOLD, L.: 'Relativistic Dynamics of a Charged Particle in Crossed Magnet and Electric Fields with Application to the Planar Magnetron', *Journal Applied Physics*, 1954, **25**, p. 683, and 1955, **26**, p. 253.
(463) GOLD, L.: 'The Relativistic Magnetron and the Effective Mass Anisotropy *Journal of Electronics*, 1956, **2**, p. 17.
(464) GOSS, T. M., and LINDSAY, P. A.: 'Spurious Modulation in Q-Band Magnetron *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 1474.
(465) GOULD, R. W.: 'Space-Charge Effects in Beam Type Magnetrons', *Journal Applied Physics*, 1957, **28**, p. 599.
(466) GUENARD, P.: 'Potential Distribution in a Cylindrical Electron Beam', *Journal Physique et le Radium*, 1945, **6**, p. 43.
(467) GUENARD, P., and HUBER, H.: 'Experimental Study of the Interaction of Spac Charge Waves within an Electron Beam moving in Crossed Electric an Magnetic Fields', *Annales de Radioélectricité*, 1952, **7**, p. 252.
(468) GUNN, J. C.: 'Theory of Radiation', *Reports on Progress in Physics*, 1955, 1 p. 127.
(469) HAGSTRUM, H. D.: 'The Generation of Centimetre Waves', *Proceedings of t Institute of Radio Engineers*, 1947, **35**, p. 548.
(470) HARKER, K. J.: 'Periodic Focusing of Beams from Partially Shielded Cathodes *Transactions of the Institute of Radio Engineers*, 1955, **ED–2**, No. 4, p. 13.
(471) HARMER, J. D.: 'A Nomogram for Hartree's Threshold-Voltage Criterion fo Magnetrons', *Electronic Engineering*, 1954, **26**, p. 441.
(472) HARRIS, L. A.: 'Axially Symmetric Electron Beam and Magnetic-Field Systems *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 700.
(473) HARRIS, L. A.: 'Instabilities in the Smooth Anode Cylindrical Magnetron *Journal of Applied Physics*, 1952, **23**, p. 562, and 1953, **24**, p. 1335.
(474) HARVEY, A. F.: 'Output and Efficiency of the Split-Anode Magnetron Oscillatin in the Dynatron Regime', *Journal I.E.E.*, 1939, **84**, p. 683.
(475) HARVEY, A. F.: 'The Cut-off Characteristic of the Single Anode Magnetron *Proceedings of the Cambridge Philosophical Society*, 1939, **35**, p. 637.
(476) HARVEY, A. F.: 'The Impedance of the Magnetron in Different Regions of th Frequency Spectrum', *Journal I.E.E.*, 1940, **86**, p. 297.
(477) HEFFNER, H., and WATKINS, D. A.: 'The Practicability of E-Type Travellin Wave Devices', *Proceedings of the Institute of Radio Engineers*, 1955, 4 p. 1007.
(478) HOK, G.: 'A Statistical Approach to the Space Charge Distribution in a Cut-o Magnetron', *Journal of Applied Physics*, 1952, **23**, p. 983.
(479) HOK, G.: 'Calculation of a Waveguide-Loaded Resonator for Interdigit Magnetrons', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 76
(480) HOLLENBERG, A. V., KROLL, N., and MILLMAN, S.: 'Rising-Sun Magnetro with Large Numbers of Cavities', *Journal of Applied Physics*, 1948, **19**, p. 62
(481) HULL, A. W.: 'The Effect of a Uniform Magnetic Field on the Motion Electrons between Coaxial Cylinders', *Physical Review*, 1921, **18**, p. 13.
(482) HULL, J. F., and RANDALS, A. W.: 'High-Power Interdigital Magnetrons *Proceedings of the Institute of Radio Engineers*, 1948, **36**, p. 1357.
(483) HULL, J. F.: 'Inverted Magnetron', *ibid.*, 1952, **40**, p. 1038.
(484) HULL, J. F., and GREENWALD, L. W.: 'Modes in Interdigital Magnetrons', *ibi* 1949, **37**, p. 1258.
(485) HUNTER, L. P.: 'Energy Build-Up in Magnetrons', *Journal of Applied Physic* 1946, **17**, p. 833, and *Physical Review*, 1946, **69**, p. 700.
(486) JELLEY, J. V.: 'Čerenkov Radiation', *Progress in Nuclear Physics*, 1953, p. 84.
(487) JEPSEN, R. L., and MULLER, M. W.: 'Enhanced Emission from Magnetro Cathodes', *Journal of Applied Physics*, 1951, **22**, p. 496.
(488) JEPSEN, R. L.: 'Thermionic and Secondary Emission Properties of Magnetro Cathodes and Their Influence on Magnetron Operation', *Physical Revie* 1950, **78**, p. 354.
(489) KALININ, V. I., and GERSTEIN, G. M.: 'On the Dynatron Effect in Multi-Segme Magnetrons', *Comptes Rendus de l'Académie des Sciences de l'URSS*, 194 **51**, p. 275.
(490) KALININ, V. I., and RYAZONOVA, T. P.: 'Theory of the Multi-Segment Ma netron', *Zhurnal Tekhnicheskoi Fiziki*, 1952, **22**, p. 1592.
(491) KIEL, A., SCOTTO, M., and PARZEN, P.: 'Propagation in a Crossed Field Period Structure', *Transactions of the Institute of Radio Engineers*, 1958, **ED–5**, No. p. 76.
(492) KLEEN, W., and POSCHL, K.: 'Focusing of Electron Beams by Magnetic Field *Archiv der Elektrischen Übertragung*, 1955, **9**, p. 295.
(493) KLEEN, W., and POSCHL, K.: 'The Generation and Amplification of Millimet Waves' *Nachrichtentechnische Zeitschrift*, 1958, **11**, pp. 8 and 77.

94) KLEIN, C. A.: 'Production of Radio Waves by means of the Čerenkov Effect', *Annales des Télécommunications*, 1953, **8**, p. 38.

95) KLEINIJEN, P. H. J. A.: 'Magnetrons', *Tijdschrift van het Nederlandsch Radiogenootschap*, 1953, **18**, p. 287.

96) KLEINWACHTER, H.: 'A Travelling-Wave Valve without Retarding Circuit', *Electrotechnische Zeitschrift*, 1951, **72**, p. 714.

97) KLEINWACHTER, H.: 'The Excitation of Electromagnetic Fields by Current Waves', *Archiv der Elektrischen Übertragung*, 1952, **6**, p. 376.

98) KOMPFNER, R.: 'Travelling-Wave Tubes', *Reports on Progress in Physics*, 1952, **15**, p. 275.

99) KROLL, N. M., and LAMB, W. E.: 'The Resonant Modes of the Rising-Sun and other Unstrapped Magnetron Anode Blocks', *Journal of Applied Physics*, 1948, **19**, p. 166.

500) LAMPERT, M. A.: 'Plasma Oscillations at Extremely High Frequencies', *ibid.*, 1956, **27**, p. 5.

501) LAMPERT, M. A.: 'Incidence of an Electromagnetic Wave on a "Čerenkov Electron Gas" ', *Physical Review*, 1956, **102**, p. 299.

502) LANDECKER, K.: 'Possibility of Frequency Multiplication and Wave Amplification by means of Relativistic Effect', *ibid.*, 1952, **86**, p. 852.

503) LANGMUIR, I.: 'The Effect of Space Charge and Residual Gases on Thermionic Currents in High Vacuum', *ibid.*, 1913, **2**, p. 450.

504) LANGMUIR, R. V., and NELSON, R. H.: 'A 10 kW Magnetron with Water-Cooled Cathode', *ibid.*, 1946, **70**, p. 118.

505) LASHINSKY, H.: 'Čerenkov Radiation from Extended Electron Beams near a Medium of Complex Index of Refraction', *Journal of Applied Physics*, 1956, **27**, p. 631.

506) LAZZERIE, M.: 'Variable Frequency Magnetrons', *L'Onde Électrique*, 1957, **37**, p. 65.

507) LEBLOND, A.: 'Investigation of an Interdigital Line used as Anode Circuit for an UHF Magnetron Oscillator: Distortions of an Electromagnetic Field', *Annales de Radioélectricité*, 1955, **10**, p. 20.

508) LEBLOND, A., NALOT, J., and DOEHLER, O.: 'Determination of the Natural Oscillations and Mode Changes of a Multicavity Magnetron Oscillator', *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences (Paris)*, 1952, **235**, p. 1495.

509) LEBLOND, A., DOEHLER, O., and WARNECKE, R.: 'New Magnetron Oscillator with Interdigital Circuit', *ibid.*, 1953, **236**, p. 55.

510) LEBLOND, A.: 'Investigation of an Interdigital Line used as Anode Circuit for a VHF Magnetron Oscillator', *Annales de Radioélectricité*, 1953, **8**, p. 194.

511) LEMMENS, H. J., JANSEN, M. J., and LOOSJES, R.: 'A New Thermionic Cathode for Heavy Loads', *Philips Technical Review*, 1950, **11**, p. 341.

512) LEVI, R.: 'New Dispenser Type Thermionic Cathode', *Journal of Applied Physics*, 1953, **24**, p. 233.

513) LEVIN, H. Y.: 'Multiple Electronic Resonance in Single Split Anode Magnetrons', *Radiotekhnika*, 1949, **4**, p. 36.

514) LINDSAY, P. A., and SIMS, G. D.: 'Application of the Thermodynamics of Irreversible Processes to the Theory of the Magnetron', *Proceedings of the Physical Society*, 1953, **66**, p. 423.

515) LINHART, J. G.: 'Čerenkov Radiation', *Research (London)*, 1955, **8**, p. 402.

516) LINHART, J. G.: 'Čerenkov Radiation of Electrons moving Parallel to a Dielectric Boundary', *Journal of Applied Physics*, 1955, **26**, p. 527.

517) LOPUKHIN, V. M.: 'The Electron Theory of a Centimetre-Wave Decelerator', *Zhurnal Tekhnicheskoi Fiziki*, 1951, **21**, p. 516.

518) LOPUKHIN, V. M.: 'The Electron Theory of the Planar Magnetron', *ibid.*, 1951, **21**, p. 505.

519) LUDI, F.: 'Theory of the Magnetron as Microwave Generator', *Helvetica Physica Acta*, 1946, **19**, p. 1.

520) LUDI, F.: 'The Turbator—A Single Cavity Magnetron', *Tijdschrift van het Nederlandsch Radiogenootschap*, 1953, **18**, p. 89.

521) LUDI, F.: 'On the Theory of the Magnetron Amplifier', *Zeitschrift für angewandte Mathematik und Physik*, 1953, **3**, p. 119.

522) MARIE, P.: 'A New Type of Magnetron Amplifier', *L'Onde Électrique*, 1950, **30**, pp. 13, 79 and 200.

523) MATHIAS, L. E. S.: 'The Space Charge Distribution in the Pre-oscillation Magnetron', *Journal of Electronics*, 1955, **1**, p. 8.

524) MATVEEV, A. N.: 'On the Optimum Length of an Undulator', *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki*, 1955, **28**, p. 6.

525) MEGAW, E. C. S.: 'The High-Power Pulsed Magnetron: a Review of Early Developments', *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 977.

526) DIX, J. C., and MEGAW, E. C. S.: 'Low-Power Resonant-Segment Magnetrons for Centimetre Waves', *ibid.*, 1946, **93**, Part IIIA, p. 1585.

527) MENDEL, J. T., QUATE, C. F., and YOCUM, W. H.: 'Electron Beam Focusing with Periodic Permanent Magnetic Fields', *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 800.

528) MENDEL, J. T.: 'Magnetic Focusing of Electron Beams', *ibid.*, 1955, **43**, p. 327.

529) MILLER, M. H., HERSHENOV, B., and BLACK, J. R.: 'Effect of Magnetic Field on Coupled Helix Attenuators', *Journal of Applied Physics*, 1957, **28**, p. 1363.

530) MILLMAN, S., and NORDSIECK, A. T.: 'The Rising Sun Magnetron', *ibid.*, 1948, **19**, p. 156.

531) MOATS, R. R.: 'Mode Interaction in Magnetrons', *Tele-Tech*, 1952, **11**, p. 39.

532) MOIZHES, B. Y.: 'On the Shape of Electron Trajectories in a Magnetron under Static Conditions', *Zhurnal Tekhnicheskoi Fiziki*, 1956, **26**, p. 1836.

533) MOTZ, H., and MALLORY, K. B.: 'Generation of Submillimetre Waves', *Journal of Applied Physics*, 1955, **26**, p. 1384.

534) MOTZ, H.: 'Čerenkov and Undulator Radiation', *Transactions of the Institute of Radio Engineers*, 1956, **AP–4**, No. 3, p. 374.

535) MOTZ, H.: 'Applications of Radiation from Fast Electron Beams', *Journal of Applied Physics*, 1951, **22**, p. 527.

536) MOTZ, H., THON, W., and WHITEHURST, R. N.: 'Experiments on Radiation by Fast Electron Beams', *ibid.*, 1953, **24**, p. 826.

537) MOURIER, G.: 'The Anticyclotron, a New Type of Travelling-Wave Valve with Magnetic Fields', *Annales de Radioélectricité*, 1950, **5**, p. 206.

538) MOURIER, G.: 'Excitation of the Carcinotron M Valve', *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences (Paris)*, 1955, **240**, p. 406.

539) NAG, B. D., and SAYIED, A. M.: 'Electrodynamics of Moving Media and the Theory of the Čerenkov Effect', *Proceedings of the Royal Society*, 1956, **235A**, p. 544.

540) NEDDERMAN, H. C.: 'Space-Charge Distribution in a Static Magnetron', *Journal of Applied Physics*, 1955, **26**, p. 1420.

541) NELSON, R. B.: 'Methods of Tuning Multiple-Cavity Magnetrons', *Physical Review*, 1946, **70**, p. 118, and *Proceedings of the Institute of Radio Engineers*, 1948, **36**, p. 53.

542) NOWOGRODZKI, M.: 'Conductance Measurement on Operating Magnetron Oscillators', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 1239.

543) OKRESS, E. C.: 'A Magnetron-Resonator System', *Journal of Applied Physics*, 1947, **18**, p. 1098.

544) OKRESS, E. C.: 'Magnetron Mode Transitions', Advances in Electronics and Electron Physics, Vol. VIII (Academic Press, New York, 1956).

545) OKRESS, E. C., GLEASON, C. H., WHITE, R. A., and HAYTER, W. R.: 'Design and Performance of a High Power Pulsed Magnetron', *Transactions of the Institute of Radio Engineers*, 1957, **ED–4**, No. 2, p. 161.

546) PAFOMOV, V. E.: 'Čerenkov Radiation in Anisotropic Ferrites', *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki*, 1956, **30**, p. 761.

547) PAFOMOV, V. E.: 'The Radiation of a Point Charge Flying Along the Boundary Separating Two Media', *ibid.*, 1957, **32**, p. 3.

548) PAGE, L., and ADAMS, N. I.: 'Space Charge n Plane Magnetron', *Physical Review*, 1946, **69**, p. 492.

549) PAGE, L., and ADAMS, N. I.: 'Space Charge in Cylindrical Magnetron', *ibid.*, 1946, **69**, p. 494.

550) PALOCZ, I.: 'The Frequency Spectrum of the Strapped Magnetron', *Acta Technica Academiae Scientiarum Hungaricae*, 1954, **9**, p. 135.

551) PANTELL, R. H., COLEMAN, P. D., and BECKER, R. C.: 'Dielectric Slow-Wave Structure for the Generation of Power at Millimetre and Submillimetre Wavelengths', *Transactions of the Institute of Radio Engineers*, 1958, **ED–5**, No. 3, p. 167.

552) PETRASCO, E., and VASILESCO, I. I.: 'A Frequency-Modulated Magnetron', *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences (Paris)*, 1957, **244**, p. 2296.

553) PHILLIPS, M., and LAMB, W. E.: 'Space Charge Frequency Dependence of a Magnetron Cavity', *Physical Review*, 1946, **69**, p. 701.

554) PIERCE, J. R.: 'The Theory and Design of Electron Beams' (Van Nostrand, New York, 1949).

555) PIERCE, J. R.: 'Spatially Alternating Magnetic Fields for Focusing Low-voltage Electron Beam', *Journal of Applied Physics*, 1953, **24**, p. 1247.

556) PIERCE, J. R., and WALKER, L. R.: 'Brillouin Flow with Thermal Velocities', *ibid.*, 1953, **24**, p. 1328.

557) POSCHL, K.: 'Theory of the Carcinotron', *Fernmeldetechnische Zeitschrift*, 1954, **7**, p. 558.

558) POSTHUMOUS, K.: 'Oscillations in a Split-Anode Magnetron', *Wireless Engineer*, 1935, **12**, p. 126.

559) POTOK, M. H. N.: 'A Critical Review of Researches into Millimetric-Wave Spark Generators', *Journal of the British Institution of Radio Engineers*, 1953, **13**, p. 490.

560) POTOK, M. H. N.: 'Researches into Spark Generation of Microwaves', *Proceedings I.E.E.*, Paper No. 2185 R, November, 1956 (**103 B**, p. 781).

561) PRAXMARER, W.: 'Calculation of the Performance Chart of Magnetrons', *Nachrichten Technische Zeitschrift*, 1956, **6**, p. 97.

562) PRAXMARER, W.: 'Design of the Anode System of Unstrapped Magnetrons for the Millimeter-Wave Range', *ibid.*, 1954, **4**, p. 159.

563) PRAXMARER, W.: 'The Wavelength Limit of the Rising Sun System and Comparison with the Unstrapped-Anode System', *ibid.*, 1954, **4**, p. 197.

564) PRAXMARER, W.: 'The Possibility of Generating Millimetre Waves with Pulsed Multislot Magnetrons of the Rising-Sun Type', *ibid.*, 1952, **2**, p. 277.

565) PURL, O. T., ANDERSON, J. R., and BREWER, G. R.: 'A High-Power Periodically Focused Traveling-Wave Tube', *Proceedings of the Institute of Radio Engineers*, 1958, **46**, p. 441.

566) RAEV, A., UZUNOV, I., and ANGELOV, A.: 'Self-Sustained Electronic Spokes in Magnetrons', *Journal of Electronics*, 1956, **1**, p. 452.

567) RANDALL, J. T., and BOOT, H. A. H.: 'Early Work on the Cavity Magnetron', *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 182.

568) RANDALL, J. T.: 'Development of the Magnetron', *Electrician*, 1946, **136**, p. 537.

569) RANDALL, J. T.: 'The Cavity Magnetron', *Proceedings of the Physical Society*, 1946, **58**, p. 247.

570) REICH, H. J., MAY, J. C., SKALNIK, J. G., and UNGAVY, R. L.: 'Some Aspects of Split-Anode Magnetron Operation', *Proceedings of the Institute of Radio Engineers*, 1950, **38**, p. 1428.

571) ROBERTS, L. W., and BRIGGS, R. S.: 'Practical Millimetre-Magnetron Considerations', *Le Vide*, 1956, **11**, p. 258, and *L'Onde Électrique*, 1957, **37**, p. 65.

572) ROBERTSHAW, R. G., and WILLSHAW, W. E.: 'Some Properties of Magnetrons using Spatial-Harmonic Operation', *Proceedings I.E.E.*, Monograph No. 168 R, February, 1956 (**103 C**, p. 297).

573) ROBERTSHAW, R. G., GOSS, T. M., TEW, J. R., and WILLSHAW, W. E.: 'A Review of the Performance of C.W. Magnetrons Operating at Low Magnetic Fields', *L'Onde Électrique*, 1957, **37**, p. 65.

574) ROGOVIN, I. E.: 'On the Possibility of Extending the Concepts of Similarity to Multiresonator Magnetrons with Different Numbers of Resonators', *Radiotekhnika i Elektronica*, 1956, **1**, p. 51.

575) SAYIED, A. M.: 'The Čerenkov Effect in Composite (Isotropic) Media', *Proceedings of the Physical Society*, 1958, **71**, p. 398.

576) SCHIFF, L. I.: 'Quantum Effects in the Radiation from Accelerated Relativistic Electrons', *American Journal of Physics*, 1952, **8**, p. 474.

577) SCHMIDT, W.: 'A.C. Operation of Magnetrons for Coherent Operation', *Elektronische Rundschau*, 1958, **12**, p. 12.

578) SENITZKY, I. R.: 'Transition from Classical to Quantum Interaction between Electron and High Frequency Fields', *Physical Review*, 1953, **90**, p. 386.

579) SENITZKY, I. R.: 'Quantum Effects in the Interaction between Electrons and High-Frequency Fields: Parts 1 and 2', *ibid.*, 1954, **95**, p. 904, and 1955, **98**, p. 875.

580) SENITZKY, I. R.: 'Quantum Effects in the Interaction between Electrons and High-frequency Fields: Vacuum Fluctuation Phenomena', *ibid.*, 1956, **104**, p. 1486.

581) SHEA, H. G.: 'Rising Sun Pulsed and C.W. Magnetrons', *Electronic Industry*, 1946, **5**, p. 46.

582) SHEVCHIK, V. N.: 'The Cascade-Bunching of Electrons in Application to the Theory of the Multi-Resonator Magnetron', *Zhurnal Tekhnicheskoi Fiziki*, 1955, **25**, p. 1462.

583) SHIMODA, K.: 'Length of Coherent Microwaves Generated by an Electron Oscillator', *Journal of the Physical Society (Japan)*, 1953, **8**, p. 131.

584) SHULMAN, C.: 'Quantum Effects in the Interaction between Free Electrons and Electromagnetic Fields', *Physical Review*, 1951, **82**, p. 116.

585) SINGH, A., and VAIDYA, N. C.: 'Fabrication of Multicavity Magnetrons', *Journal of Scientific and Industrial Research*, 1957, **16A**, p. 169.

586) SINGH, A.: 'Modes and Operating Voltages of Interdigital Magnetrons', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 470.

587) SIMS, G. D., and GABOR, D.: 'Theory of the Pre-Oscillating Magnetron: Part 2—Perturbations of a Double Stream Steady State', *Journal of Electronics*, 1955, **1**, p. 231.

588) SIRKIS, M. D., and COLEMAN, P. D.: 'The Harmodotron—a Megavolt Electronics Millimetre Wave Generator', *Journal of Applied Physics*, 1957, **28**, p. 944.

589) SLATER, J. C.: 'Microwave Electronics' (Van Nostrand, New York, 1950).

590) SMITH, L. P.: 'Quantum Effects in the Interaction of Electrons with High Frequency Fields and the Transition to Classical Theory', *Physical Review*, 1946, **69**, p. 195.

591) SOLOMON, S. S.: 'Space-Charge Waves in Crossed Electric and Magnetic Fields', *Journal of Applied Physics*, 1955, **26**, p. 1443.

592) SUSSKIND, C.: 'Electron Guns and Focusing for High Density Electron Beams', Advances in Electronics and Electron Physics. Volume VIII (Academic Press, New York, 1956).

(593) SVENSON, R.: 'An Experimental Investigation of the Electron Orbits in a Magnetron', *Proceedings of the Institute of Radio Engineers*, 1951, **39**, p. 838.
(594) SWANN, W. F. G.: 'Certain Matters Concerning Scaling in the Magnetron with Special Reference to the Relative Efficiency of Magnetrons of Different Sizes', *Journal of the Franklin Institute*, 1948, **246**, p. 149.
(595) TCHERNOV, Z. S.: 'The Spiratron—A Centrifugal Electrostatic Focusing Travelling-Wave Tube', *L'Onde Électrique*, 1957, **37**, p. 65.
(596) TIEN, P. K.: 'Focusing of a Long Cylindrical Electron Stream by means of Periodic Electrostatic Fields', *Journal of Applied Physics*, 1954, **25**, p. 1281.
(597) TOPPINGA, M. L., and SCHUYTEMAKER, J.: 'Problems in the Coupling of the Magnetron to a Waveguide', *Tijdschrift van het Nederlandsch Radiogenootschap*, 1954, **19**, p. 157.
(598) TWISS, R. Q.: 'On the Initial Space-Charge Distribution in a Cylindrical Magnetron Diode', *Journal of Electronics*, 1955, **1**, p. 1.
(599) TYCHINSKI, M., and DERKACH, Y. T.: 'Oscillations of the Space Charge Cloud in a Cylindrical Magnetron', *Radiotekhnika i Elektronika*, 1956, **1**, pp. 233 and 344.
(600) VALLANTIN, B.: 'Asymmetries in Strapped Resonator Systems of Magnetrons', *Le Vide*, 1956, **11**, p. 243, and *L'Onde Électrique*, 1957, **37**, p. 65.
(601) VAUGHAN, J. R. M.: 'A Millimetre-Wave Magnetron', *Proceedings I.E.E.*, Monograph No. 142 R, August, 1955 (103 C, p. 95).
(602) VAUGHAN, J. R. M.: 'An Experimental Cold-Cathode Magnetron', *Le Vide*, 1956, **11**, p. 251, and *L'Onde Électrique*, 1957, **37**, p. 65.
(603) VERSNEL, I. A.: 'Magnetless Magnetrons', *L'Onde Électrique*, 19 **7**, **37**, p. 65, and *Le Vide*, 1957, **12**, p. 59.
(604) VERWEEL, J.: 'The Space Charge Distribution in a Magnetron under Static Conditions', *Le Vide*, 1957, **12**, p. 32.
(605) VERWEEL, J.: 'Magnetrons', *Philips Technical Review*, 1952, **14**, p. 44.
(606) VOGE, J.: 'The Internal Mechanism of the Magnetron', *L'Onde Électrique*, 1946, **26**, pp. 345 and 374.
(607) WALL, T. F.: 'The Magnetron and the Klystron', *Engineering*, 1946, **161**, pp. 125, 148 and 184.
(608) WANG, C. C.: 'Electron Beams in Axially Symmetrical Electric and Magnetic Fields', *Proceedings of the Institute of Radio Engineers*, 1950, **38**, p. 135.
(609) WARD, J. C.: 'Quantum Effects in the Interaction of Electrons with High-Frequency Fields', *Physical Review*, 1950, **80**, p. 119.
(610) WARNECKE, R., KLEEN, W., LERBS, A., DOEHLER, O., and HUBER, H.: 'The Magnetron-Type Traveling-Wave Amplifier Tube', *Proceedings of the Institute of Radio Engineers*, 1950, **38**, p. 486.
(611) WARNECKE, R., DOEHLER, O., and BOBOT, D.: 'Space-Charge Effects in Magnetic-Field Travelling Wave Valves', *Annales de Radioélectricité*, 1950, **5**, p. 279.
(612) WARNECKE, R., HUBER, H., GUENARD, P., and DOEHLER, O.: 'Amplification by Space-Charge Waves in an Electron Beam acted on by Crossed Electric and Magnetic Fields', *Comptes Rendus* (*Paris*), 1952, **235**, p. 470.
(613) WARNECKE, R., and GUENARD, P.: 'Some Recent Work in France on New Types of Valves for the Highest Radio Frequencies', *Proceedings I.E.E.*, 1953, **100**, Part III, p. 351.
(614) WARNECKE, R., GUENARD, P., DOEHLER, O., and EPSZTEIN, B.: 'The M Type Carcinotron Tube', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 413.
(615) WARNECKE, R.: 'The Evolution of the Principles of Modern Electronic Tubes for Microwaves', Congegno di Electronica e Televisione (Milan), April, 1954.
(616) WARNECKE, R., NALOT, J., EPSZTEIN, B., and DOEHLER, O.: 'The Carmatron, a New Electronically Tuned Wide Band Oscillator', *Comptes Rendus* (*Paris*), 1955, **241**, p. 695.
(617) WARNECKE, R., and DOEHLER, O.: 'On the Interaction Between an Electromagnetic Wave and an Electron Beam Travelling in a Cylindrical System Perpendicular to an Electric Field with Crossed Magnetic Field', *ibid.*, 1950, 231, p. 1132.
(618) WARNECKE, R., DOEHLER, O., and KLEEN, W.: 'Amplification of Electromagnetic Waves by Interaction between Electrons Moving in Crossed Electric and Magnetic Fields', *ibid.*, 1949, **229**, p. 709.
(619) WATKINS, D. A., and WADA, G.: 'The Helitron Oscillator', *Proceedings of the Institute of Radio Engineers*, 1958, **46**, p. 1700.
(620) WELCH, H. W.: 'Effects of Space Charge on Frequency Characteristics of Magnetrons', *ibid.*, 1950, **38**, p. 1434.
(621) WELCH, H. W.: 'Prediction of Travelling-Wave-Magnetron Frequency Characteristics: Frequency Pushing and Voltage Tuning', *ibid.*, 1953, **41**, p. 1631.
(622) WELCH, H. W., and DOW, W. G.: 'Analysis of Synchronous Conditions in the Cylindrical Magnetron Space Charge', *Journal of Applied Physics*, 1951, **22**, p. 433.
(623) WILBUR, D. A.: 'New Magnetron Designs for Continuous Operation in the Decimetre Wave Range', *Physical Review*, 1946, **70**, p. 118.
(624) WILLSHAW, W. E., and RUSHFORTH, L.: 'The High Power Pulsed Magnetron: An Outline of the Mechanism of Operation', *Journal I.E.E.*, 1946, 93, Part IIIA, p. 180.
(625) WILLSHAW, W. E., RUSHFORTH, L., STAINSBY, A. G., LATHAM, R., BALLS, A. W., and KING, A. H.: 'The High-Power Pulsed Magnetron: Development and Design for Radar Application', *ibid.*, 1946, **93**, Part IIIA, p. 985.
(626) WILLSHAW, W. E., and ROBERTSHAW, R. G.: 'The Behaviour of Multiple-Circuit Magnetrons in the Neighbourhood of the Critical Anode Voltage', *Proceedings of the Physical Society*, 1950, **63**, p. 41.
(627) WILLSHAW, W., MOURIER, G., and GUILBAUD, G.: 'Electronic Resonance Effect in Valves with Crossed Electric and Magnetic Fields', *Comptes Rendus* (*Paris*), 1955, **240**, p. 283.
(628) YOUNG, R. T., HALMBAR, L. W., and WATERS, W. E.: 'Some Observations on the Back Heating of Magnetron Cathodes', *Journal of Applied Physics*, 1950, 22, p. 1066.
(629) ZAMOROZHOV, B. M.: 'Phase Focusing in a Magnetron with Plane Electrodes', *Zhurnal Tekhnicheskoi Fiziki*, 1949, 19, p. 1321.
(630) ZWOBADA, R.: 'High Power Millimetric Magnetron', *L'Onde Électrique*, 1957, 37, p. 65.

Part 3

(631) AGDUR, B. N.: 'Experimental Investigation of Noise Reduction in Travelling-Wave Tubes', *Chalmers Tekniska Hogskolas Handlingar*, 1954, No. 139, p. 1.
(632) ALLEN, D. H. O., and WINWOOD, J. M.: 'A Low-Noise Travelling-Wave-Tube Amplifier for the 4 000 Mc/s Communications Band', *Journal of the British Institution of Radio Engineers*, 1957, **17**, p. 75.
(633) BEAM, W. R.: 'Interception Noise in Electron Beams at Microwave Frequencies'. *RCA Review*, 1955, **16**, p. 551.
(634) BEAM, W. R.: 'Noise Wave Excitation at the Cathode of a Microwave Beam Amplifier', *Transactions of the Institute of Radio Engineers*, 1957, **ED-4**, No. 3, p. 226.
(635) BELL, R. L.: 'Klystron Oscillator Noise Theory', *British Journal of Applied Physics*, 1956, **7**, p. 262.
(636) BLOOM, S., and PETER, R. W.: 'A Minimum Noise Figure for the Traveling-Wave Tube', *RCA Review*, 1954, **15**, p. 252.
(637) BLOOM, S.: 'The Effect of Initial Noise Current and Velocity Correlation on Noise Figure of Traveling-Wave Tubes', *ibid.*, 1955, **16**, p. 179.
(638) BOYD, J. A.: 'Noise Characteristics of a Voltage Tuneable Magnetron', *Transactions of the Institute of Radio Engineers*, 1954, **ED-1**, No. 4, p. 201.
(639) BUCHMILLER, L. D., DEGRASSE, R. W., and WADE, G.: 'Design and Calculation Procedures for Low-Noise Traveling-Wave Tubes', *ibid.*, 1957, **ED-4**, No. p. 234.
(640) BURKE, P. F. C., and POHL, W. J.: 'A 4 000 Mc/s Low-Noise Travelling-Wave Tube', *L'Onde Électrique*, 1957, **37**, p. 65.
(641) CAULTON, M., and ST. JOHN, G. E.: 'S-Band Traveling-Wave Tube with Noise Figure below 4 dB', *Proceedings of the Institute of Radio Engineers*, 1958, p. 911.
(642) CONVERT, G.: 'Experimental Low-Noise Amplifying Valve', *Annales de Radioélectricité*, 1952, **7**, p. 225.
(643) CONVERT, G., and DOEHLER, O.: 'The Signal to Noise Ratio in the M-Carcinotron', *Transactions of the Institute of Radio Engineers*, 1954, **ED-1**, No. p. 184.
(644) CURRIE, M. R., and FORSTER, D. C.: 'Low-Noise Tuneable Preamplifiers for Microwave Receivers', *Proceedings of the Institute of Radio Engineers*, 1 **46**, p. 570.
(645) CURRIE, M. R., and FORSTER, D. C.: 'Condition for Minimum Noise Generated in Backward-Wave Amplifiers', *Transactions of the Institute of Radio Engine* 1958, **ED-5**, No. 2, p. 88.
(646) CURRIE, M. R., and FORSTER, D. C.: 'Experiments on Noise Reduction Backward-Wave Amplifiers', *Proceedings of the Institute of Radio Engine* 1957, **45**, p. 690.
(647) CUTLER, C. C., and QUATE, C. F.: 'Verification of Transit Time Reduction Noise', *Physical Review*, 1950, **80**, p. 875.
(648) ESPERSEN, G. A.: 'A Low-Noise High-Power Klystron Oscillator of G Reliability', *Le Vide*, 1956, **11**, p. 270, and *L'Onde Électrique*, 1957, p. 65.
(649) EVERHART, T. E.: 'Concerning the Noise Figure of a Backward-Wave Amplifier', *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 444.
(650) HARRISON, S. W.: 'A Low-Noise Traveling-Wave Tube', *Sylvania Technolog* 1950, **3**, p. 12.
(651) HARRISON, S. W.: 'On the Minimum Noise Figure of Traveling-Wave Tub *Proceedings of the Institute of Radio Engineers*, 1955, **43**, p. 227, and *Sylv Technologist*, 1954, **7**, p. 123.
(652) HAUS, H. A.: 'Noise in One-Dimensional Electron Beam', *Journal of App Physics*, 1955, **26**, p. 560.
(653) HAUS, H. A., and ROBINSON, F. N. H.: 'The Minimum Noise Figure of Mi wave Beam Amplifiers', *Proceedings of the Institute of Radio Engineers*, 1 **43**, pp. 981 and 1101.
(654) HAUS, H. A.: 'Limitations on the Noise Figure of Microwave Amplifiers the Beam Type', *Transactions of the Institute of Radio Engineers*, 1 **ED-1**, No. 4, p. 238.
(655) HOK, G.: 'The Noise Factor in Traveling-Wave Tubes', *Proceedings of Institute of Radio Engineers*, 1956, **44**, p. 1061.
(656) KINAMAN, E. W., and MAGID, M.: 'Very-Low-Noise Travelling-Wave Amplif *ibid.*, 1958, **46**, p. 861, and *Le Vide*, 1957, **12**, p. 308.
(657) KLEEN, W.: 'On the Noise in Transit Time Tubes', *Frequenz*, 1952, **6**, p. 45.
(658) KLEEN, W., and RUPPEL, W.: 'Calculation of the Noise Figure of the Travel Wave Valve', Parts 1 and 2, *Archiv der Elektrischen Übertragung*, 1952 pp. 187 and 299.
(659) KNECHTLI, R. C.: 'Effect of Electron Lenses on Beam Noise', *Transaction the Institute of Radio Engineers*, 1958, **ED-5**, No. 2, p. 84.
(660) KNECHTLI, R. C., and BEAM, W. R.: 'Validity of Traveling-Wave-Tube N Theory', *RCA Review*, 1957, **18**, p. 24.
(661) KNECHTLI, R. C., and BEAM, W. R.: 'Design and Performance of Low-N Guns for Traveling-Wave Tubes', Convention Record of the Institute Radio Engineers, 1956, Part 3, p. 23, and *RCA Review*, 1956, **17**, p. 410.
(662) LABUS, J., LIEBSCHER, R., and POSCHL, M.: 'Conditions for the Minim Noise Figure of Travelling-Wave Valves', *Archiv der Elektrischen Übertrag* 1956, **10**, p. 486.
(663) LAPLANTE, R. A.: 'Development of a Low Noise X-Band C.W. Klystron Po Oscillator', *Transactions of the Institute of Radio Engineers*, 1954, **ED-1**, No p. 99.
(664) LEHR, C. G., and COLLINS, A. L.: 'Physical Mechanism of Noise Genera in Magnetrons', *ibid.*, 1954, **ED-1**, No. 4, p. 260.
(665) LESOTA, S. K.: 'Minimum Noise Coefficient of Double-Stream Valve', *Ra tekhnika i Elektronika*, 1956, **1**, p. 1288.
(666) MIDDLETON, D., GOTTSCHALK, W. M., and WIESNER, J. B.: 'Noise in C Magnetrons', *Journal of Applied Physics*, 1953, **24**, p. 1065.
(667) MUNGALL, A. G.: 'Noise in Traveling Wave Tubes', *Transactions of the Inst of Radio Engineers*, 1955, **ED-2**, No. 2, p. 12.
(668) PARZEN, P.: 'Effect of Thermal-Velocity Spread on the Noise Figure Travelling-Wave Tubes', *Journal of Applied Physics*, 1952, **23**, p. 394.
(669) PEIFOR, A. G., PARZEN, P., and BRYANT, J. H.: 'Low-Noise Traveling-W Tube', *Electrical Communication*, 1952, **29**, p. 234.
(670) PETER, R. W.: 'Low-Noise Traveling-Wave Amplifier', *RCA Review*, 1 **13**, p. 344.
(671) PETER, R. W., and RUETZ, J. A.: 'Influence of Secondary Electrons on N Factor and Stability of Traveling-Wave Tubes', *ibid.*, 1953, **14**, p. 441.
(672) PETERSON, L. C.: 'Space-Charge and Transit-Time Effects on Signal and N in Microwave Tetrodes', *Proceedings of the Institute of Radio Engineers*, 1 **35**, p. 1264.
(673) PIERCE, J. R.: 'A New Method of Calculating Microwave Noise in Elect Streams', *ibid.*, 1952, **40**, p. 1675.
(674) PIERCE, J. R.: 'A Theory Concerning Noise in Electron Streams', *Journa Applied Physics*, 1954, **25**, p. 931.
(675) PIERCE, J. R., and DANIELSON, J. R.: 'Minimum Noise Figure of Travell Wave Tubes with Uniform Helices', *ibid.*, 1954, **25**, p. 1163.
(676) PIERCE, J. R.: 'The General Sources of Noise in Vacuum Tubes', *Transact of the Institute of Radio Engineers*, 1957, **ED-1**, No. 4, p. 135.
(677) RIGROD, W. W.: 'Noise Spectrum of Electron Beam in Longitudinal Magn Field', *Bell System Technical Journal*, 1957, **36**, pp. 831 and 855.
(678) ROBINSON, F. N. H., and KOMPFNER, R.: 'Noise in Traveling Wave Tub *Proceedings of the Institute of Radio Engineers*, 1951, **39**, p. 918.
(679) ROBINSON, F. N. H.: 'Microwave Shot Noise in Electron Beams and Minimum Noise Factor of Travelling Wave Tubes and Klystrons', *Journa the British Institution of Radio Engineers*, 1954, **14**, p. 79.
(680) ROBINSON, F. N. H., and HAUS, H. A.: 'Analysis of Noise in Electron Bea *Journal of Electronics*, 1956, **1**, p. 373.
(681) ROBINSON, F. N. H.: 'Space Charge Smoothing of Microwave Shot No *Philosophical Magazine*, 1952, **63**, p. 51.
(682) ROBINSON, F. N. H.: 'Microwave Shot Noise and Amplifiers', *Transaction the Institute of Radio Engineers*, 1956, **ED-3**, p. 128.
(683 SCHNITZER, H., and WEBER, D.: 'A Method for the Reduction of the N Factor of Travelling-Wave Valves', *Fernmeldetechnische Zeitschrift*, 195 p. 302.

(34) SCHNITZER, H., and WEBER, D.: 'Low-Noise Travelling Wave Valves', *ibid.*, 1954, 7, p. 540.
(35) SIEGMAN, A. E., and WATKINS, D. A.: 'Potential Minimum Noise in the Microwave Diode', *Transactions of the Institute of Radio Engineers*, 1957, ED–4, No. 1, p. 82.
(36) SIEGMAN, A. E., and BLOOM, S.: 'An Equivalent Circuit for Microwave Noise at the Potential Minimum', *ibid.*, 1957, ED–4, No. 4, p. 295.
(37) SIEGMAN, A. E., WATKINS, D. A., and HSIEH, H. C.: 'Density-Function Calculations of Noise Propagation on an Accelerated Multivelocity Beam', *Journal of Applied Physics*, 1957, 28, p. 1138.
(38) STEWART, J. L.: 'Theory of Frequency Modulation Noise in Tubes Employing Phase Focusing', *ibid.*, 1955, 26, p. 409.
(39) STRUM, P. D.: 'Klystron Noise', *Transactions of the Institute of Radio Engineers*, 1955, MTT–3, No. 1, p. 45.
(40) TAGER, A. S.: 'Investigation of Noise Characteristics of Travelling-Wave Valves', *Radiotekhnika i Elektronika*, 1957, 2, p. 222, *Le Vide*, 1956, 11, p. 379, and *L'Onde Électrique*, 1957, 37, p. 65.
(691) TIEN, P. K.: 'A Dip in the Minimum Noise Figure of Beam-Type Microwave Amplifier', *Proceedings of the Institute of Radio Engineers*, 1956, 44, p. 938.
(692) TIEN, P. K., and MOSHMAN, J.: 'Monte Carlo Calculation of Noise Near the Potential Minimum of a High Frequency Diode', *Journal of Applied Physics*, 1956, 27, p. 1067.
(693) WADE, G., AMO, K., and WATKINS, D. A.: 'Noise in Transverse-Field Travelling-Wave Tubes', *ibid.*, 1954, 25, p. 1514.
(694) WATKINS, D. A.: 'Traveling-Wave Tube Noise Figure', *Proceedings of the Institute of Radio Engineers*, 1952, 40, p. 65.
(695) WATKINS, D. A.: 'Low-Noise Traveling Wave Tubes for X-Band', *ibid.*, 1953, 41, p. 1741.
(696) WATKINS, D. A.: 'Noise at the Potential Minimum in the High Frequency Diode', *Journal of Applied Physics*, 1955, 26, p. 622.
(697) WHINNERY, J. R.: 'Noise Phenomena in the Region of the Potential Minimum', *Transactions of the Institute of Radio Engineers*, 1954, ED–1, No. 4, p. 221.
(698) YADAVILLI, S. V.: 'Tube Noise under Large Transit-Time Conditions', *Journal of Applied Physics*, 1954, 25, p. 564.

DISCUSSION ON THE ABOVE MONOGRAPH

Dr. W. A. Gambling and **Mr. B. G. Bosch** (*communicated*): In such an extensive review as Dr. Harvey's it is perhaps a little unfair to criticize any one part. However we feel that Section 8 is a little short and slightly misleading. For example, in tubes employing long electron beams an appreciable amount of noise may also be caused by the presence of positive ions. Sometimes these manifest themselves as discrete oscillations which may be easily observed, but often they cause an increase of the noise sidebands over a range of frequencies. Another point is that modulation noise is not necessarily restricted to a narrow bandwidth, but can occur over quite a broad band.

No mention is made in the paper of noise in O-type backward-wave oscillators, and it may therefore be of interest to quote some typical measurements which we have made at X-band frequencies in order to illustrate the above comments. A useful figure of merit for some applications is the r.m.s. carrier frequency deviation, and for a C.S.F. CO43 O-type carcinotron under normal operating conditions the value obtained is 6·4 kc/s, with a correlation coefficient between the f.m. and a.m. components of 0·35. For comparison, the corresponding figures for a CV323 klystron are 6·9 kc/s and 0·57. The bandwidth of the system used in making these measurements was about 4 Mc/s.

Noise sidebands of the same carcinotron and, again for comparison, of a Varian klystron are shown in Figs. A and B. The

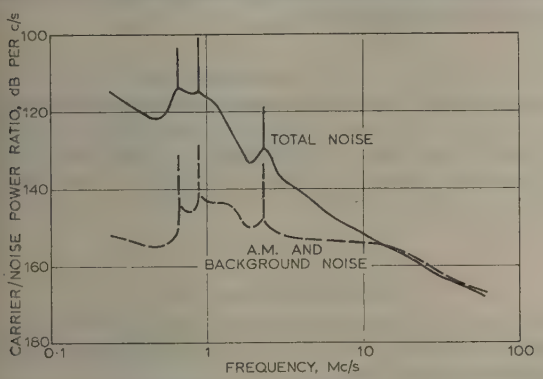


Fig. A.—Total noise and a.m. plus background noise sidebands of O-type carcinotron CO43 at an operating frequency of 9·3 Gc/s.

upper curve in Fig. A gives the total (i.e. sum of a.m., f.m. and background) noise, while the lower curve shows the a.m. plus background component only. The difference between these is a measure of the f.m. noise. The noise peaks at 0·67, 0·88 and 2·3 Mc/s are typical effects due to the presence of positive ions, and the general increase in noise level over a broad frequency band is clearly shown. It may also be seen not only that f.m. noise exists near the carrier frequency, as implied in the paper, but that, in, fact it predominates at sideband frequencies up to

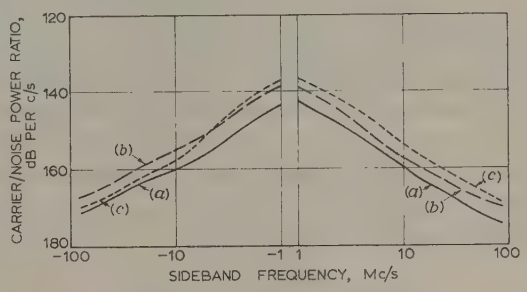


Fig. B.—Upper and lower total noise sidebands of reflex klystron V58 at an operating frequency of 9·3 Gc/s.
(*a*) A mode centre.
(*b*) At upper half-power point.
(*c*) At lower half-power point.

12 Mc/s. A separate measurement at higher frequencies has not yet been made, but it is hoped to do this shortly. With local-oscillator tubes only the a.m. plus background noise is of interest, but with transmitting tubes it is the total noise which is of importance. For the klystron the upper and lower sidebands were measured separately, both at the mode centre and at the half-power points. As might be expected the carrier/noise power ratio is best at the mode centre. The results obtained with adjacent modes differed by only one or two decibels, the noise level increasing with mode number. It is interesting to note that in the range 20–60 Mc/s the noise levels of the klystron and the carcinotron are similar, although the power supplies for the latter were highly stabilized.

To the already impressive bibliography we should like to add two more references concerning noise in backward-wave oscillators, which appeared before May, 1958.

REFERENCES

(A) KRULEE, R. L.: 'Carcinotron Noise Measurements,' *Transactions of the Institute of Radio Engineers*, 1954, **ED1**, No. 4. p. 131.
(B) CICCHETTI, J. B., and MINUSHIAN, J.: 'Noise Characteristics of a Backward-Wave Oscillator', *Institute of Radio Engineers National Convention Record*, 1958, Part 3, p. 84.

Dr. A. F. Harvey (*in reply*): The term 'narrow band' was employed in the paper in a relative sense to distinguish noise due to modulation causes from that due to shot and similar effects which, making due allowance for any filtering action of the output microwave structure, is sensibly constant over an extremely wide range of frequency. Section 8 is necessarily brief because little has been published on modulation noise in electron tubes, and in this respect the contribution of Dr. Gambling and Mr. Bosch, reporting observations of effects due to ion oscillations and measurements extending beyond 10 Mc/s, is most welcome.

621.372.54: 621.391

The Institution of Electrical Engin
Monograph No. 3
Nov. 1

THE OPTIMIZATION OF A CLASS OF NON-LINEAR FILTERS

By J. K. LUBBOCK, M.A., B.Sc.(Eng.), Graduate.

(*The paper was first received* 23*rd February, and in revised form* 2*nd July*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in November*, 1959.)

SUMMARY

An optimum linear filter, in the mean-square-error sense, is no better than the optimum attenuator if there is no dissimilarity between the spectral densities of the signal and the noise. However, non-linear filters can use more statistical information about the signal and noise, so that, although they both possess the same spectral densities, a non-linear filter may be able to introduce a significant improvement in the mean-square error (i.e. it can do better than the optimum attenuator). For this reason the use of a non-linear filter in certain circumstances may well justify the greater difficulties encountered in its optimization and physical realization.

The class of filters considered in the paper may be defined by a general expression relating output to input:

$$y(t) = \sum_{r=1}^{R} \int_0^\infty \omega_r(\tau)\theta_r[x(t-\tau)]d\tau$$

An almost routine procedure is proposed whereby the optimum set of weighting functions, $\omega_r(\tau)$, can be determined given either long enough samples of the combined input and the signal or sufficient statistical information about their characteristics. Some worked examples demonstrate that:

(*a*) A significant improvement in mean-square error is possible even under the condition when the signal and noise possess the same spectral densities.

(*b*) The class of non-linear filters under consideration can be optimized unhindered by the need to evaluate difficult integrals.

(*c*) Although the physical complexity of the filter increases rapidly with the value of R, the mean-square error may converge rapidly to an asymptotic value as R is increased; in one example the performance was found to be within about 5% of the asymptotic value with $R = 2$.

(1) INTRODUCTION

In communication and control systems it often happens that information transmitted by the system becomes corrupted by the addition of an unwanted component called noise. The noise may be due to a number of causes such as crosstalk, distortion or random disturbances. The relative magnitude of this effect depends on the system and its input. In some cases it is quite unimportant, in others the noise and signal may be so inextricably combined that the original signal information is veritably 'lost'. As a measure of this effect we shall use the mean-square error:

$$\overline{e^2} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T}[y(t) - z(t)]^2 dt = \overline{[y(t) - z(t)]^2} \quad . \quad (1)$$

where $y(t)$ is the system output and $z(t)$ is the desired output, e.g. the signal component of input.

If the input to a system is already corrupted by noise, it may be possible, by correct design, to arrange for the mean-square error at its output to be less than that at its input. Such a system is called a noise filter. It may be necessary for a system to predict future values of the signal from a knowledge of its past values, in which case it is called a predictor. The use of linear devices for these purposes is well known, but, as sho in the paper, non-linear devices can often give greater reduct in mean-square error.

A filtering device must be able to identify and use some similarity between the statistical character of the noise signal components of its input. If the statistical descriptions the signal and noise characteristics are identical in all respe then no significant improvement in mean-square error is possi i.e. the optimum* filter becomes an attenuator. Furtherm if we choose to optimize a filter subject to the restrict that it belongs to a specified class of filters, we also rest the kind of information which the filter can identify use. Thus an optimum filter belonging to the linear class filters can produce a significant improvement in mean-squ error, only if the spectral densities of the signal and the n are different, since it cannot use any other information. O dissimilarities in the characteristics of the signal and noise wh the linear filter has ignored, e.g. amplitude dissimilarities, be used by non-linear filters; thus a non-linear filter may able to produce a significant improvement in mean-square er although a linear filter cannot do so.

(2) SOME PHYSICALLY REALIZABLE FILTERS

The following is a classification of some simple filters wh can be physically realized and optimized using the theoret method outlined in the paper.

(2.1) Attenuator

An attenuator is a device† whose output at some instan equal to the sum of a term which is proportional to the inpu the same instant and a d.c. level term. (By this definition allow that the constant of proportionality may be greater t unity.)

i.e. $$y(t) = kx(t) + K \quad . \quad . \quad . \quad . \quad .$$

(2.2) Zero-Memory Filter

The output of a zero-memory filter is given by an instantane function of the input. The class of attenuators forms a sub-cl of the zero-memory class.

$$y(t) = \mathrm{f}[x(t)] = \int_0^\infty \mathrm{f}[x(t-\tau)]\delta(\tau)d\tau \quad . \quad . \quad .$$

where $\delta(\tau)$ is the Dirac delta function.

(2.3) Linear Filter

The output of a linear filter, in general, can be expressed

* We define 'optimum filter' as the filter which gives the lowest value of mean-sq error at its output subject to the usual restriction that it belongs to a specified of filters.

† Although mathematically an attenuator is a special type of filter, and it improve the mean-square error, any reduction in mean-square error that it produce is not considered to be significant for reasons explained in Appendix 8.3.

Correspondence on Monographs is invited for consideration with a view to publication.

Mr. Lubbock is in the Department of Engineering, University of Cambridge.

continuous linear summation of all past values of input in he form of the convolution integral

$$y(t) = \int_0^\infty x(t-\tau)\omega(\tau)d\tau \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where $\omega(t)$ is the filter-weighting-function or delta-function esponse and $x(t)$ is the filter input. The class of attenuators orms a sub-class of the linear class.

(2.4) Some Non-Linear Memory Filters

A simple non-linear memory filter can be realized by connecting a zero-memory filter and a linear filter in cascade. Expres-

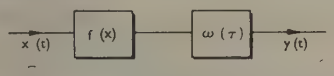


Fig. 1.—Zero-memory filter followed by linear memory filter.

ions relating output to input are given below for several anonical forms of memory filters.

$$y(t) = \int_0^\infty \mathrm{f}[x(t-\tau)]\omega(\tau)d\tau \quad . \quad . \quad . \quad . \quad (5)$$

see Fig. 1)

$$y(t) = \sum_{r=0}^{R} \int_0^\infty \mathrm{f}_r[x(t-\tau)]\omega_r(\tau)d\tau \quad . \quad . \quad . \quad (6)$$

Both eqns. (5) and (6) are special cases of the following equation which defines a class of filters, η_1:

$$y(t) = \int_0^\infty \mathrm{K}[x(t-\tau), \tau]d\tau \quad . \quad . \quad . \quad . \quad (7)$$

Expanding the kernel function $\mathrm{K}[x, \tau]$ in terms of a set of inearly independent or orthogonal functions $\theta_r(\tau)$,

$$y(t) = \sum_{r=0}^{\infty} \int_0^\infty \theta_r[x(t-\tau)]\omega_r(\tau)d\tau \quad . \quad . \quad . \quad (8)$$

By using a finite number of terms in the expansion, the filter can be physically realized (Fig. 2) and made to approximate to

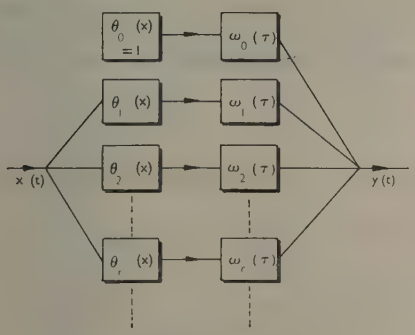


Fig. 2.—Multi-path memory filter.

any filter defined by eqn. (7) with any desired accuracy, assuming that the expansion converges. Filters for which the input and output are related by an expression of the form of eqn. (7) are members of class η_1. The classes of linear and zero-memory filters are both sub-classes of η_1, since eqns. (3) and (4) are of the same form as eqn. (7). A filter of class η_1 which is realized using only R parallel paths, each path representing a term of the series expansion, is called a multi-path filter of class η_{1R}. The class η_{1R} is a sub-class of η_{1N} when $N > R$ and η_{1N} is a sub-class of η_1.

Still more general classes of non-linear memory filters can be defined[1, 2] by generalizing eqn. (7) and in other ways.[3, 4]

(3) OPTIMIZATION OF FILTERS OF CLASS η_1

The optimization procedure outlined below applies not only to filters for separating signal from noise with which it has become contaminated but also to the more general case in which the desired output depends on the input or some part of it in some linear or non-linear manner, as exemplified by a predictor or a demodulator. The term 'filter' is used in the general sense to cover all such cases, and, where necessary, the term 'noise filter' is used for the special application of separating signal from noise.

The optimization procedure consists of minimizing the mean-square error defined by eqn. (1):

$$\overline{e^2} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} [y(t) - z(t)]^2 dt \quad (1)$$

When the filter is optimized as a noise filter, the desired output, $z(t)$, is equal to the signal component of input, and

$$x(t) = \mathrm{g}[z(t), v(t)] \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where $v(t)$ is the noise component of the input and g is a function of both z and v. Two examples will be mentioned:

(*a*) In transmission of intelligence (by line, radio or other means) noise may be added linearly to the signal (e.g. atmospheric noise, thermal noise, shot noise, crosstalk, etc.); thus

$$x(t) = z(t) + v(t) \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where usually $z(t)$ and $v(t)$ are statistically independent.

(*b*) Also in transmission, a signal can become distorted due to non-linearities of the transmission system; in this case noise is added in the form of distortion components; thus

$$x(t) = z(t) + \mathrm{h}[z(t)] \quad . \quad . \quad . \quad . \quad . \quad (11)$$

where h is a function of z, often instantaneous, containing no linear terms. In this case $z(t)$ and $\mathrm{h}[z(t)]$ usually possess some correlation.

When the filter is optimized as a predictor $z(t)$ is the desired output of the predictor and is normally expressed as a function of $x(t + T)$:

$$z(t) = \mathrm{f}[x(t+T)] \quad . \quad . \quad . \quad . \quad . \quad (12)$$

or, more commonly,

$$z(t) = x(t+T) \quad . \quad . \quad . \quad . \quad . \quad (13)$$

For the class of filters η_1,

$$y(t) = \int_0^\infty \mathrm{K}[x(t-\tau), \tau]d\tau \quad . \quad . \quad . \quad . \quad (7)$$

and the mean-square error is

$$\overline{e^2} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \left\{\int_0^\infty \mathrm{K}[x(t-\tau), \tau]d\tau - z(t)\right\}^2 dt$$

Changing the order of integration and expanding the integrand,

$$\overline{e^2} = \int_0^\infty \int_0^\infty \overline{\mathrm{K}[x(t-\tau_1), \tau_1]\mathrm{K}[x(t-\tau_2), \tau_2]}d\tau_1 d\tau_2 - 2\int_0^\infty \overline{\mathrm{K}[x(t-\tau), \tau]z(t)}d\tau + \overline{z(t)^2} \quad . \quad (14)$$

If $x(t)$ and $z(t)$ are members of ensembles $X(t)$ and $Z(t)$ which are both ergodic, time averages may be replaced by ensemble averages, giving:*

$$\overline{e^2} = \int_0^\infty \int_0^\infty \int \int K(x_1, \tau_1)K(x_2, \tau_2)p_2(x_1, x_2; \tau_2 - \tau_1)dx_1dx_2d\tau_1d\tau_2 - 2\int_0^\infty \int \int K(x, \tau)zp_2(x, z; \tau)dxdzd\tau + \overline{z^2} \quad (15)$$

where $p_2(x_1, x_2; \tau)$ is the second probability density of the process $X(t)$, and $p_2(x, z; \tau)$ is the joint second probability density of the processes $X(t)$ and $Z(t)$. The minimization of this expression is achieved by using variational calculus.

Let $K(x, \tau)$ be changed by an amount $\epsilon H(x, \tau)$, where ϵ is a small parameter and $H(x, \tau)$ is an arbitrary kernel function subject only to the restriction that it must represent a physically realizable filter, i.e. $H(x, \tau) = 0$ $\tau < 0$.

The variation in $\overline{e^2}$ is given by†

$$\Delta\overline{e^2} = 2\epsilon\int_0^\infty \int_0^\infty \int \int K(x_2, \tau_2)H(x_1, \tau_1)p_2(x_1, x_2; \tau_1 - \tau_2)dx_1dx_2d\tau_1d\tau_2 - 2\epsilon\int_0^\infty \int \int H(x, \tau)zp_2(x, z; \tau)dxdzd\tau + \epsilon^2\int_0^\infty \int_0^\infty \int \int H(x_1, \tau_1)H(x_2, \tau_2)p_2(x_1, x_2; \tau_1 - \tau_2)dx_1dx_2d\tau_1d\tau_2 \quad . \quad . \quad . \quad . \quad (16)$$

By putting $\partial\Delta\overline{e^2}/\partial\epsilon = 0$ when $\epsilon = 0$,

$$\int_0^\infty \int H(x_1, \tau_1)\left[\int_0^\infty \int K(x_2, \tau_2)p_2(x_1, x_2; \tau_1 - \tau_2)dx_2d\tau_2 - \int zp_2(x_1, z; \tau_1)dz\right]dx_1d\tau_1 = 0 \quad . \quad (17)$$

Also from eqn. (16) it can be shown that $\partial^2\Delta\overline{e^2}/\partial\epsilon^2$ is always positive, so that the stationary value of $\overline{e^2}$ is a minimum. Using the fundamental lemma of variational calculus,[5] we obtain from eqn. (17) an integral equation which gives the implicit solution for the kernel function of the optimum non-linear filter:

$$\int_0^\infty \int K(x_2, \tau_2)p_2(x_1, x_2; \tau_1 - \tau_2)dx_2d\tau_2 = \int zp_2(x_1, z; \tau_1)dz \quad . \quad . \quad . \quad . \quad (18)$$

which must be satisfied for $\tau_1 \geqslant 0$ only, since $H(x_1, \tau_1) = 0 \tau_1 < 0$. Thus the information which is required in order to optimize a filter of class η_1 is the second probability density of the input $p_2(x_1, x_2; \tau_1 - \tau_2)$ and the joint second probability density $p_2(x_1, z; \tau_1)$ between the input and desired output.

(3.1) Zero-Memory Filters

The input-output relationship for a zero-memory filter is of the form

$$y(t) = f[x(t)] = \int_0^\infty f[x(t - \tau)]\delta(\tau)d\tau \quad . \quad . \quad . \quad (3)$$

Eqn. (3) shows that zero-memory filters are members of t more general class η_1 and the kernel function in eqn. (7) is the form

$$K[x(t - \tau), \tau] = f[x(t - \tau)]\delta(\tau)$$

Substitution in eqn. (17) for the kernel functions $K(x_2, \tau_2)$ $f(x_2)\delta(\tau_2)$ and $H(x_1, \tau_1) = h(x_1)\delta(\tau_1)$ and subsequent integrati with respect to τ_1 and τ_2 leads to the equation

$$\int h(x_1)\left[\int f(x_2)p_2(x_1, x_2; 0)dx_2 - \int zp_2(x_1, z; 0)dz\right]dx_1 = 0$$

whence, using the lemma of variational calculus,

$$\int f(x_2)p_2(x_1, x_2; 0)dx_2 = \int zp_2(x_1, z; 0)dz \quad . \quad . \quad (1$$

But $p_2(x_1, x_2; 0) = p(x_1)\delta(x_2 - x_1)$

Therefore eqn. (19) can be rewritten

$$f(x)p(x) = \int zp_2(x, z)dz$$

giving an explicit solution for $f(x)$, namely

$$f(x) = \frac{\int zp_2(x, z)dz}{p(x)} \quad . \quad . \quad . \quad . \quad . \quad (2$$

where $p_2(x, z)\delta x\delta z$ is the probability that x lies in the range to $x + \delta x$ and z lies in the range z to $z + \delta z$ simultaneously.

Thus the optimum filter function $f(x)$ is equal to the condition mean value of z, given x. The information needed in t optimization of a zero-memory filter is the whole informati content of the first probability density function of the input a the second joint probability density function of the input a desired output with $\tau = 0$.

(3.2) Multi-Path Zero-Memory Filters

For many processes it is possible to obtain an expansion f $p_2(x, z; \tau)$ in the form [see Appendix 8.1, eqn. (58)]

$$p_2(x, z; \tau) = \sum_{n=0}^\infty \sum_{m=0}^\infty D_{nm}(\tau)\theta_n(x)\gamma_m(z)p(x)q(z) \quad . \quad (5$$

where $\theta_n(x)$ and $\gamma_m(z)$ are two sets of polynomials which a orthonormal with respect to $p(x)$ and $q(z)$ as weighting functio over the whole range of x and z, respectively, and $p(x)$ and q are first probability density functions of x and z, respectivel also

$$D_{nm}(\tau) = \iint \theta_n(x)\gamma_m(z)p_2(x, z; \tau)dxdz \quad . \quad . \quad (5$$

whence, by putting $\tau = 0$ and substituting in eqn. (20),

$$f(x) = \sum_{n=0}^\infty \sum_{m=0}^\infty D_{nm}\theta_n(x) \int z\gamma_m(z)q(z)dz$$

But $$\int z\gamma_m(z)q(z)dz = \begin{cases} 0 & m \neq 0; m \neq 1 \\ \sigma z & m = 1 \\ \bar{z} & m = 0 \end{cases}$$

owing to the orthogonality of $\gamma_m(z)$.

Therefore $$f(x) = \sum_{n=1}^\infty D_{n1}\theta_n(x)\sigma_z + \bar{z} \quad . \quad . \quad . \quad . \quad (2$$

since, as shown in Appendix 8.1, $D_{00} = 1$, $\theta_0(x) = 1$ a $D_{0n} = D_{n0} = 0$, $n \neq 0$.

* Throughout the paper, (i) for each integral written without limits, it is implied that integration takes place over the whole range of the variable of integration, and (ii) all processes are assumed to be ergodic.

† The first term in the expression for $\Delta\overline{e^2}$ is the sum of two terms which can be combined, since $p_2(x_1, x_2; \tau_2 - \tau_1)$ is symmetrical in the pairs of variables (x_1, τ_1) and (x_2, τ_2), i.e. $p_2(x_1, x_2; \tau_2 - \tau_1) = p_2(x_2, x_1; \tau_1 - \tau_2)$.

It is shown in Appendix 8.3 that truncation of eqn. (21) at the Rth term yields the optimum zero-memory multi-path filter of degree R. Therefore it can be synthesized approximately as shown in Fig. 3.

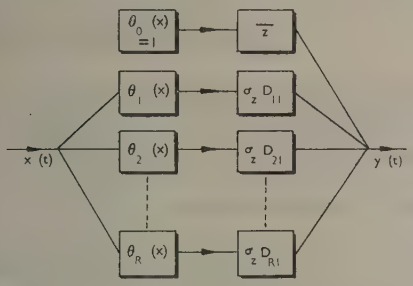


Fig. 3.—Multi-path zero-memory filter of degree R.

The following recurrence relations, derived in Appendix 8.2, show that the polynomials $\theta_n(x)$ can be determined from a knowledge of the set of moments of x up to and including the $2n$th, and the coefficients $\sigma z D_{n1}$ may be determined from a knowledge of the moments of x and the averaged cross-products between z and x^r for $r = 1, 2 \ldots n$:

$$\theta_n(x) = \frac{x^n - \sum_{r=0}^{n-1} \overline{x^n\theta_r(x)}\theta_r(x)}{\left[\overline{x^{2n}} - \sum_{r=0}^{n-1} \overline{x^n\theta_r(x)}^2\right]^{1/2}} \quad . \quad . \quad . \quad (80)$$

and, from eqns. (79), (81) and (82), putting $\tau = 0$,

$$\sigma_z D_{n1} = \overline{\theta_n(x)z} = \frac{\overline{x^n z} - \sum_{r=0}^{n-1} \overline{x^n\theta_r(x)}D_{r1}\sigma_z}{\left[\overline{x^{2n}} - \sum_{r=0}^{n-1} \overline{x^n\theta_r(x)}^2\right]^{1/2}} \quad . \quad . \quad (22)$$

A finite number of coefficients D_{n1} [in eqn. (21)] and constants σ_z and $\bar{z}$ can be determined experimentally by analogue measurements on long samples of the input, $x(t)$, and desired output, $z(t)$. The required computing equipment consists of:

(*a*) A set of non-linear power-law-function generators, namely

$$x, x^2, x^3, x^4, \ldots \quad \ldots, x^{2n}$$

(A convenient device which will generate all of these functions simultaneously is a servo-type multiplier with $(2n + 1)$ potentiometers connected to its shaft.)

(*b*) An integrator.
(*c*) A set of variable attenuators.
(*d*) A set of adders.
(*e*) A multiplier.

The procedure is as follows:

(i) Measure the set of moments (up to the $2n$th) of the input process, $X(t)$, by connecting the signal, $x(t)$, to the inputs of the function generators and averaging their outputs in turn using the integrator.

(ii) Determine the appropriate orthonormal polynomials up to the nth using the recurrence equation [eqn. (80)].

(iii) Measure the set of averaged cross-products $\overline{x^r z}$ by averaging over time using the power-law generators, the multiplier and integrator; then compute the coefficients from the recurrence equation [eqn. (22)].

Alternatively, set up the first n polynomial-function generators, $\theta_n(x)$, using the power-law generators, the attenuators and adders and measure the coefficients $\sigma_z D_{n1}$ directly using the multiplier and integrator. For an ergodic process we can express $\sigma_z D_{n1}$ as a time average:

$$\sigma_z D_{n1} = \int\int \theta_n(x) z \mathrm{p}_2(x, z; 0)dxdz$$

$$= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \theta_n[x(t)]z(t)dt$$

A block diagram of the measuring circuit is shown in Fig. 4.

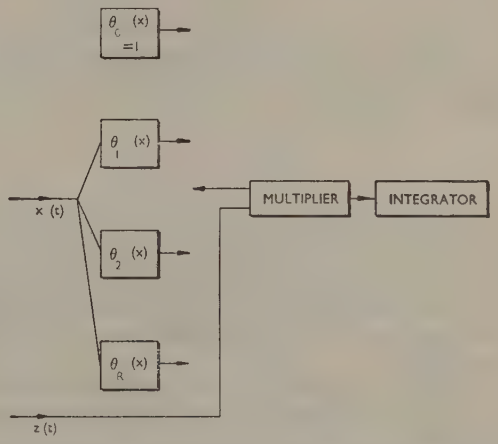


Fig. 4.—Experimental method of measuring the constants D_{n1}.

(3.3) Multi-Path Non-Linear Memory Filters

The expression for the output of the multi-path filter of Fig. 2 can be obtained by substituting in eqn. (7) an expansion for the kernel function:

$$\mathrm{K}[x(t-\tau), \tau] = \sum_{r=0}^{\infty} \omega_r(\tau)\theta_r[x(t-\tau)] \quad . \quad . \quad (23)$$

Thus by substituting for the kernel function

$$\mathrm{K}[x_2, \tau_2] = \sum_{r=0}^{\infty} \omega_r(\tau_2)\theta_r(x_2) \quad . \quad . \quad . \quad . \quad (24)$$

in eqn. (18) for the optimum non-linear filter of class η_1, and at the same time substituting series expansions for $\mathrm{p}_2(x_1, x_2; \tau)$ and $\mathrm{p}_2(x_1, z; \tau)$ (see Appendix 8.1 and 8.2), namely

$$\mathrm{p}_2(x_1, x_2; \tau) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} C_{nm}(\tau)\theta_n(x_1)\theta_m(x_2)\mathrm{p}(x_1)\mathrm{p}(x_2) \quad . \quad (83)$$

$$\mathrm{p}_2(x_1, z; \tau) = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} D_{kl}(\tau)\theta_k(x_1)\gamma_l(z)\mathrm{p}(x_1)\mathrm{q}(z) \quad . \quad (58)$$

we obtain

$$\sum_{r=0}^{\infty}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \theta_n(x_1)\mathrm{p}(x_1) \times$$

$$\int_0^{\infty} C_{nm}[(\tau_1 - \tau_2)]\omega_r(\tau_2)d\tau_2 \int \theta_m(x_2)\theta_r(x_2)\mathrm{p}(x_2)dx_2$$

$$= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} D_{kl}(\tau_1)\theta_k(x_1)\mathrm{p}(x_1)\int z\gamma_l(z)\mathrm{q}(z)dz \quad . \quad (25)$$

$$\tau_1 \geqslant 0$$

But $$\int \theta_m(x_2)\theta_r(x_2)p(x_2)dx_2 = 1 \quad m = r$$
$$= 0 \quad m \neq r$$

and $$\int z\gamma_l(z)q(z)dz = 0 \quad l \neq 0;\ l \neq 1$$
$$= \sigma_z \quad l = 1$$
$$= \bar{z} \quad l = 0$$

Using these results and equating coefficients of $\theta_n(x_1)$ for $0 \leqslant n \leqslant R$ in eqn. (25),

$$\sum_{m=1}^{R} \int_0^\infty C_{nm}[(\tau_1 - \tau_2)]\omega_m(\tau_2)d\tau_2 = D_{n1}(\tau_1)\sigma z \quad . \quad (26)$$
$$\tau_1 \geqslant 0 \qquad n = 1, 2, 3 \ldots R$$

and when $m = n = 0$

$$\int_0^\infty \omega_0(\tau_2)d\tau_2 = \bar{z} \quad . \quad . \quad . \quad . \quad . \quad (27)$$

In general, the optimum kernel function can be expanded only as an infinite expansion [eqn. (24)], but in order that a solution for the optimum weighting functions $\omega_m(\tau)$ can be found, truncation of eqn. (24) is necessary, thereby yielding a finite number, R, of simultaneous integral equations [eqn. (26)], in which, in general, the weighting functions $\omega_m(\tau)$ for $m \leqslant R$ will depend on R.

It is shown in Appendix 8.3 that the weighting functions given by eqns. (26) are those of the optimum multi-path memory filter of degree R.

The simultaneous integral equations [eqn. (26)] can be solved by the method of undetermined coefficients[6] (see Appendix 8.4). It seems that often there is a unique set of solutions, provided that additional restrictions on the weighting functions are imposed to ensure physical realizability in a sufficiently strict sense (see Appendix 8.4). However, if a set of solutions to eqns. (26) is not unique, then, at least, it can be shown to give a mean-square error which is no greater than for any other set (Appendix 8.3).

The multi-path filter of class η_{1R} can only approximate to a general filter of class η_1, because for physical realizability the number of paths, R, must be finite. For a large value of R more information is required to optimize the filter until in the limit as R tends to infinity the whole information content of $p_2(x_1, x_2; \tau)$ and a large part of the information in $p_2(x, z; \tau)$ is required, i.e. all the coefficients $C_{nm}(\tau)$ in the infinite expansion of $p_2(x_1, x_2; \tau)$ and $D_{n1}(\tau)$ in the infinite expansion of $p_2(x, z; \tau)$.

The coefficients $C_{nm}(\tau)$ and $\sigma_z D_{n1}(\tau)$ can be determined analytically as described in Appendix 8.2 given sufficient statistical information concerning the input and desired output, or they can be determined experimentally by making analogue measurements on long samples of the input, $x(t)$, and desired output, $z(t)$. The required computing equipment is the same as that listed in Section 3.2 except that a cross-correlator is required in place of the multiplier. A very suitable cross-correlator for this purpose is described by Lampard.[7] The procedure for making the measurements is as follows:

(*a*) Measure the set of moments of the input signal, $x(t)$, up to and including the $2n$th moment, using the power-law-function generators and the integrator.

(*b*) Determine the appropriate orthonormal polynomials up to the nth using eqn. (80).

(*c*) Set up the first n polynomial-function generators, $\theta_n(x)$, using the power-law generators, the attenuators and adders.

(*d*) Measure the auto- and cross-correlation functions between the outputs of the function generators taking two at a time. These functions will be the coefficients $C_{nm}(\tau)$, since, fr eqn. (84) for an ergodic process,

$$C_{nm}(\tau) = \iint \theta_n(x_1)\theta_m(x_2)p_{xx}(x_1, x_2; \tau)dx_1dx_2$$
$$= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \theta_n[x(t)]\theta_m[x(t+\tau)]dt \quad . \quad ($$

(*e*) Measure the cross-correlation function between each o put of the function generators and the desired filter output, z These correlation functions will be the coefficients $\sigma_z D_{n1}$ since, from eqn. (59) for an ergodic process,

$$\sigma_z D_{n1}(\tau) = \iint \theta_n(x)zp_{xz}(x, z; \tau)dxdz$$
$$= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \theta_n[x(t)]z(t+\tau)dt \quad . \quad (n \neq 0$$

The measuring circuit is shown in Fig. 5.

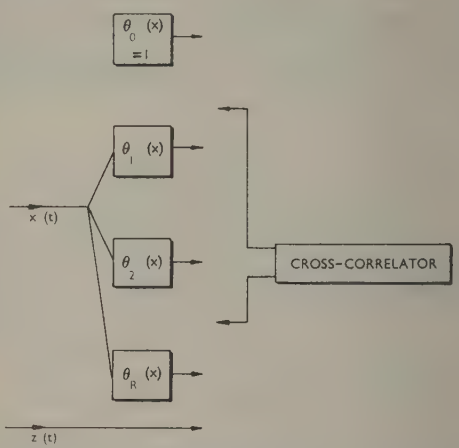


Fig. 5.—Experimental method of measuring the coefficients C_{nm} and $D_{n1}(\tau)$.

Although for many cases of input and desired output a n linear filter may be capable of giving considerably smaller me square filtering error than can an optimum linear filter, there situations in which the performance of an optimum non-lin filter is no better than the optimum linear filter.

If* $C_{nm}(\tau) \equiv 0$ when $n = 1 \quad m \neq 1$
and $m = 1 \quad n \neq 1$

and $D_{nm}(\tau) \equiv 0$ when $m = 1 \quad n \neq 1$
$\tau \geqslant 0$ } . . (

eqn. (26) can be satisfied by putting

$$\omega_m(\tau) \equiv 0 \qquad m = 2, 3, \ldots \infty$$

and $$\int_0^\infty C_{11}(\tau_1 - \tau_2)\omega_1(\tau_2)d\tau_2 = D_{11}(\tau_1)\sigma_z \quad . \quad .$$
$$\tau, \geqslant 0$$

Therefore $\omega_1(\tau_2)$ is given by the solution of the Wiener–H integral equation [eqn. (31)], in which

$$C_{11}(\tau) = \rho_{xx}(\tau)$$
$$D_{11}(\tau) = \rho_{xz}(\tau) \quad \} \quad . \quad . \quad . \quad . \quad . \quad .$$

* These conditions define a class of probability density functions $p_2(x_1, x_2; \tau)$ $p_2(x, z, \tau)$. This class is not empty and some of its properties are discusse References 8, 9 and 10.

re the normalized correlation coefficients. Although we do not rove that this solution is unique, it does give a value of mean-quare error which is as small as for any other solution Appendix 8.3). Thus the optimum filter of class η_1 can be a inear filter, and the only information which is needed for its ptimization is the auto-correlation of the input and the cross-orrelation of the input and desired output.

(4) SOME NOISE-FILTER EXAMPLES

Two examples are worked out to illustrate the method. In he first example the signal and noise can have the same spectral ensity, but since only a zero-memory filter is optimized the pectral density (or auto-correlation coefficient) is not required; or the case in question it is found possible to determine the xact zero-memory-filter function from eqn. (20), and for com-arison the multi-path zero memory filter is also derived, so hat some indication of the degree of approximation is obtained. n the second example, a multi-path memory filter is optimized or the condition when the signal and noise have the same pectral density, a significant improvement in mean-square error eing achieved.

Whilst examples are worked out from statistical data, it may appen that not enough information about the statistical proper-ies of the signal and noise is available to enable the filter to be ptimized; however, if long enough samples of signal and noise re available, then, using the analogue measuring technique escribed earlier, sufficient information can be obtained in a uitable form to allow direct optimization of a non-linear filter f class η_{1R}. Whilst there may be practical difficulties involved n obtaining sufficient accuracy of measurement, it is considered hat a computer of comparable accuracy to that of Pace* should e adequate to enable these analogue measurements to be nade.

Under certain circumstances it may be possible to find a impler and better non-linear filter belonging to some class ther than η_1, in which event it would obviously not be esirable to use the optimization procedure proposed here. lthough the proposed method does not always produce the implest solution, it can be used for almost any condition of nput and desired output and will always produce a physically ealizable filter which is better than the best linear filter, except n the special circumstances [eqn. (30)] when the optimum filter f class η_1 becomes linear.

In all the following examples we shall optimize no more than he first five paths of the multi-path filter.

For the first example we take an input consisting of a linear ddition of signal and noise, where the signal is a random quare wave which can have two values $+K$ or $-K$, each occur-ing with equal probability (i.e. the signal has zero mean and ariance K^2).

$$p_z(z) = \tfrac{1}{2}[\delta(z - K) + \delta(z + K)] \quad . \quad . \quad . \quad (33)$$

The noise is Gaussian with variance σ^2 and zero mean.

$$p_v(v) = \frac{1}{\sqrt{(2\pi)\sigma}}\varepsilon^{-\frac{v^2}{2\sigma^2}} \quad . \quad . \quad . \quad . \quad (34)$$

The optimum zero-memory-filter function is found first and hen the multi-path zero-memory filter is optimized and the esults are compared.

Since the noise and signal are independent and added linearly,

$$p_2(x, z) = p_v(x - z)p_z(z)$$

$$= \frac{1}{2\sqrt{(2\pi)\sigma}}[\delta(z - K) + \delta(z + K)]\varepsilon^{-\frac{(x-z)^2}{2\sigma^2}}$$

* Precision Analogue Computing Equipment.

The optimum zero-memory filter is given by eqn. (20); hence

$$f(x) = \frac{\int_{-K}^{K} \frac{1}{2\sqrt{(2\pi)\sigma}} z[\delta(z - K) + \delta(z + K)]\varepsilon^{-\frac{(x-z)^2}{2\sigma^2}} dz}{\int_{-K}^{K} \frac{1}{2\sqrt{(2\pi)\sigma}}[\delta(z - K) + \delta(z + K)]\varepsilon^{-\frac{(x-z)^2}{2\sigma^2}} dz}$$

$$= K\frac{\varepsilon^{-\frac{(x-K)^2}{2\sigma^2}} - \varepsilon^{-\frac{(x+K)^2}{2\sigma^2}}}{\varepsilon^{-\frac{(x-K)^2}{2\sigma^2}} + \varepsilon^{-\frac{(x+K)^2}{2\sigma^2}}}$$

$$= K \tanh \frac{Kx}{\sigma^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (35)$$

As K/σ increases the error decreases. For small values of K/σ there is a high probability that x will be negative for a positive value of z, such a condition giving rise to large error (see Fig. 6).

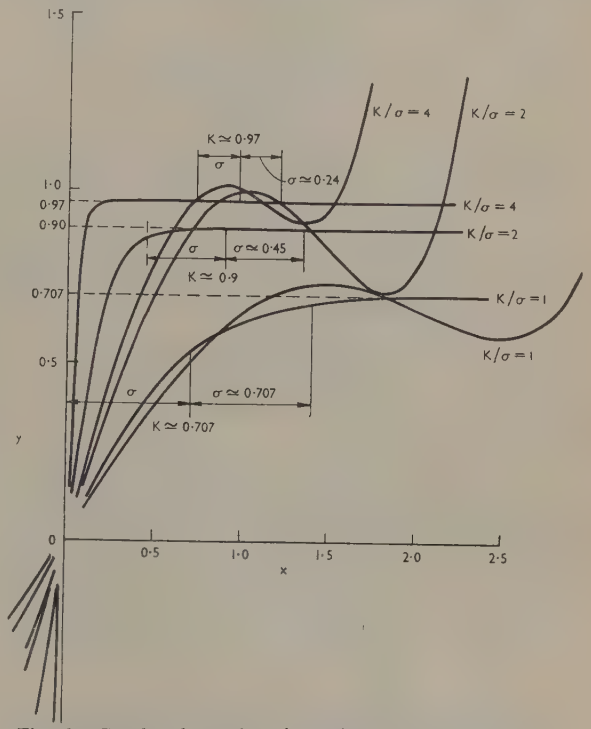


Fig. 6.—Graphs of $K \tanh Kx/\sigma_2$ and the optimum multi-path zero-memory-filter function for three different values of K/σ. $K^2 + \sigma^2 = 1$

The multi-path filter may be optimized as follows:

The first moments of a Gaussian signal with zero mean and variance σ^2 are given by

$$\left.\begin{aligned} \overline{v^n} &= \sigma^n \times 1 \times 3 \times 5 \times 7 \times \ldots (n-1) \quad n \text{ even} \\ \overline{v^n} &= 0 \qquad n \text{ odd} \end{aligned}\right\} \quad . \quad (36)$$

The first moments of a random square wave with amplitude K are $\overline{z^n} = K^n$ for n even and $\overline{z^n} = 0$ for n odd. For simplicity we will let $K^2 = \sigma^2 = \frac{1}{2}$ so that the variance of their sum is unity. Using eqn. (73) (Appendix 8.2) the moments are found from $x(t) = z(t) + v(t)$ to be

$$\overline{x^2} = 1 \quad \overline{x^4} = 2{\cdot}5 \quad \overline{x^6} = 9{\cdot}5 \quad \overline{x^8} = 47{\cdot}75 \quad \overline{x^{10}} = 296{\cdot}75$$
$$. \quad . \quad . \quad . \quad (37)$$

whereas odd moments are zero.

Using the recurrence formula 80 [eqn. (80), Appendix 8.2] the polynomials are found to be

$$\left.\begin{aligned}
\theta_0(x) &= 1 \\
\theta_1(x) &= x \\
\theta_2(x) &= \frac{x^2 - 1}{\left(\frac{3}{2}\right)^{1/2}} \\
\theta_3(x) &= \frac{x^3 - 2{\cdot}5x}{\left(\frac{13}{4}\right)^{1/2}} \\
\theta_4(x) &= \frac{x^4 - \frac{14}{3}x^2 + \frac{13}{6}}{\left(\frac{53}{6}\right)^{1/2}} \\
\theta_5(x) &= \frac{x^5 - \frac{96}{13}x + \frac{233}{26}x}{\left(\frac{1\,522}{52}\right)^{1/2}}
\end{aligned}\right\} \quad . \quad . \quad . \quad (38)$$

The coefficients D_{n1} are then found using eqn. (81) in Appendix 8.2 for $n = 1$, 3 and 5:

$$D_{11} = \frac{1}{\sqrt{2}} \qquad D_{31} = -\frac{1}{\sqrt{26}} \qquad D_{51} = \frac{9}{\sqrt{13}\sqrt{761}} \quad . \quad (39)$$

To compare this result with that obtained earlier [eqn. (35)] the polynomial represented by the expression

$$y = [D_{11}\theta_1(x) + D_{31}\theta_3(x) + D_{51}\theta_5(x)]\sigma_z$$

is worked out and plotted on the same axes as the curve $y = K \tanh Kx/\sigma^2$ when $K = 1/\sqrt{2}$ and $\sigma^2 = \frac{1}{2}$ (see Fig. 6). For these values of K and σ^2,

$$y = 0{\cdot}798\,3x - 0{\cdot}164\,3x^3 + 0{\cdot}011\,83x^5 \quad . \quad . \quad (40)$$

Since in this case $K/\sigma\,(= 1)$ is relatively small, the optimum zero-memory filter is far from perfect. The performance index (Appendix 8.3) has been worked out for the multi-path zero-memory filter for 1, 3 and 5 paths (Fig. 7) using eqns. (105) and (113).

The curve of Fig. 7 suggests that an increase in the number of paths would not reduce the performance index below abo 0·90 when $K/\sigma = 1$. Fig. 7 also shows the performance ind plotted for two other values of K/σ. The polynomial filt functions for these cases are given below and plotted in Fig.

$$K/\sigma = 2 \qquad y = 1{\cdot}553\,0x - 0{\cdot}641\,1x^3 + 0{\cdot}087\,59x^5 \quad . \quad (4$$

$$K/\sigma = 4 \qquad y = 1{\cdot}837\,7x - 1{\cdot}071\,6x^3 + 0{\cdot}237\,8x^5 \quad . \quad (4$$

It is to be expected that very much better performance w be obtained for larger values of K/σ.

This first example is almost trivial because the exact optimu zero-memory filter can be found using eqn. (20), and hen there is no advantage in optimizing the multi-path zero-memo filter in this case. The problem has been worked out primar to give some indication of the nearness of the approximati (see Fig. 6) obtained by using only a few paths.

For the second case of signal and noise the optimum mul path filters, both with and without memory, will be four Consider a noise $v(t)$ which is Gaussian (of zero mean and u variance) and a signal which is obtained by passing Gaussi noise through a non-linear-function generator of the form

$$z(t) = \frac{H_2[s(t)]}{\sqrt{2}} = \frac{s^2 - 1}{\sqrt{2}} \quad . \quad . \quad . \quad . \quad (12$$

where $s(t)$ is Gaussian noise of zero mean and unit variance b uncorrelated with $v(t)$, and H_2 is the second-order Herm polynomial.

From eqn. (125) it can be shown that $z(t)$ has zero mean a unit variance and that if the correlation coefficient of $s(t)$ exponential so also is that of $z(t)$.* The reason for choosi this method of generating the signal lies in the lat property of non-linear Hermite-polynomial generators wi respect to a Gaussian input. Since the correlation coefficie of $z(t)$ is both known and simple, we can choose the cor lation coefficient of $v(t)$ to be equal to that of $z(t)$, with t object of making a comparison between the optimum non-line filter and the optimum linear filter, which for this case becom a simple attenuator. Unfortunately, optimization of a ze memory filter using eqn. (20) leads to integrals which are n easy to evaluate (except by numerical methods); however, t difficulty does not prohibit us from finding the optimum mul path zero-memory filter by the method outlined in Appendix 8

The moments of Gaussian noise, $v(t)$, with zero-mean and u variance, are

$$\left.\begin{aligned}
\overline{v^n} &= 1 \times 3 \times 5 \times \ldots (n-1) \quad n \text{ even} \\
&= 0 \quad n \text{ odd}
\end{aligned}\right\} \quad . \quad . \quad (3$$

and the moments of $z(t)$ (as calculated in Appendix 8.5) are

$$\left.\begin{aligned}
&\overline{z} = 0 \quad \overline{z^2} = 1 \quad \overline{z^3} = 2{\cdot}828\,427 \quad \overline{z^4} = 15 \\
&\overline{z^5} = 96{\cdot}166\,522 \quad \overline{z^6} = 755 \quad \overline{z^7} = 6\,983{\cdot}386\,571 \\
&\overline{z^8} = 74\,417 \quad \overline{z^9} = 897\,799{\cdot}337\,9 \quad \overline{z^{10}} = 12\,096\,873
\end{aligned}\right\} \quad . \quad (4$$

The moments of $x(t)$ are obtained, given†

$$x(t) = \frac{v(t) + z(t)}{\sqrt{2}} \quad . \quad . \quad . \quad . \quad . \quad (4$$

by using eqn. (73), Appendix 8.2:

$$\left.\begin{aligned}
&\overline{x} = 0 \quad \overline{x^2} = 1 \quad \overline{x^3} = 1 \quad \overline{x^4} = 6 \quad \overline{x^5} = 22 \quad \overline{x^6} = 130 \\
&\overline{x^7} = 822 \quad \overline{x^8} = 6\,202 \quad \overline{x^9} = 52\,552 \quad \overline{x^{10}} = 499\,194
\end{aligned}\right\} \quad . \quad (4$$

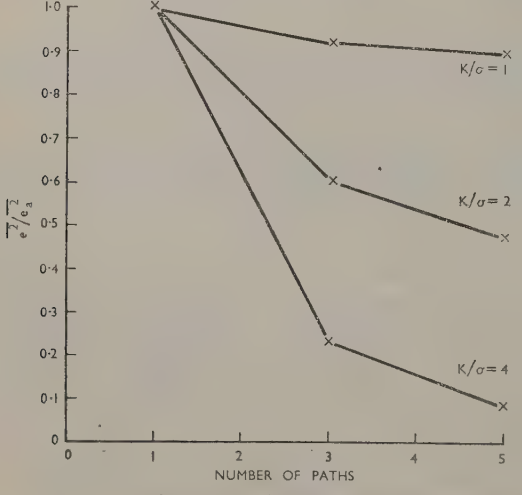


Fig. 7.—Performance index for the first problem.

* Both the second probability density functions of the signal and noise are of c Λ. (See Appendix 8.1 and 8.5, and Reference 11.)

† $\sqrt{2}$ is introduced here for convenience so that $\overline{x^2} = 1$.

The polynomials may then be constructed using the recurrence formula [eqn. (80), Appendix 8.2] giving

$$\left.\begin{aligned}
\theta_0(x) &= 1\\
\theta_1(x) &= x\\
\theta_2(x) &= \tfrac{1}{2}(x^2 - x - 1)\\
\theta_3(x) &= \frac{1}{6\cdot062\,2}(x^3 - 3\cdot75x^2 - 2\cdot25x + 2\cdot75)\\
\theta_4(x) &= \frac{1}{24\cdot646\,9}(x^4 - 8\cdot204\,1x^3 + 5\cdot265\,3x^2\\
&\qquad + 21\cdot959\,2x - 3\cdot061\,2)\\
\theta_5(x) &= \frac{1}{125\cdot075\,5}(x^5 - 14\cdot462\,8x^4 + 40\cdot082\,6x^3\\
&\qquad + 50\cdot991\,7x^2 - 103\cdot305\,8x - 26\cdot297\,5)
\end{aligned}\right\} \quad . \quad (46)$$

The coefficients, D_{n1}, are obtained using the recurrence formula [eqn. (81), Appendix 8.2]:

$$\left.\begin{aligned}
D_{11} &= 0\cdot707\,1\\
D_{21} &= 0\cdot353\,6\\
D_{31} &= -0\cdot087\,5\\
D_{41} &= -0\cdot038\,6\\
D_{51} &= 0\cdot061\,7
\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (47)$$

The final polynomial then becomes

$$y = D_{11}\theta_1(x) + D_{21}\theta_2(x) + \ldots D_{51}\theta_5(x)$$

Therefore $y = 0\cdot000\,5x^5 - 0\cdot008\,7x^4 + 0\cdot018\,2x^3$
$+ 0\cdot247\,8x^2 + 0\cdot477\,4x - 0\cdot224\,6 \quad . \quad (48)$

This function has been plotted in Fig. 8. The optimum multi-path memory filter can be found by calculating the set of cross- and auto-correlation coefficients between pairs of all the outputs of the non-linear function generators, $\theta_n(x)$, and the desired output, $z(t)$. The method outlined in Appendix 8.2, although of routine nature, does take a considerable time. The results are quoted here for a maximum of only three polynomial generators.*

where the input auto-correlation coefficient $\rho(\tau) = \varepsilon^{-\beta|\tau|}$. Eqns. (26) are then solved by the method of undetermined coefficients (Appendix 8.4) to give the weighting functions $\omega_n(\tau)$.

For only two parallel paths (it so happens that the two integral equations are independent in this case),

$$\left.\begin{aligned}
\omega_1(\tau) &= 0\cdot707\,1\delta(\tau)\\
\omega_2(\tau) &= 0\cdot276\,1\delta(\tau) + 0\cdot222\,2\beta\varepsilon^{-1\cdot1952\beta\tau} \quad \tau \geqslant 0\\
&= 0 \quad \tau < 0
\end{aligned}\right\} \quad . \quad (51)$$

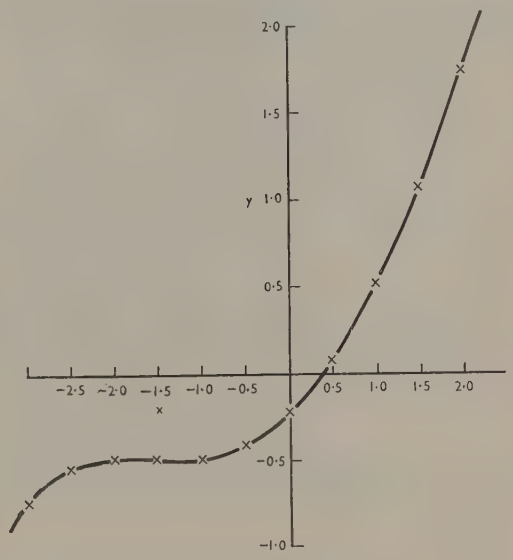


Fig. 8.—Zero-memory-filter function for the second problem.

$$\left.\begin{aligned}
C_{11}(\tau) &= \rho(\tau) \quad C_{1n}(\tau) \equiv C_{n1}(\tau) \equiv 0\\
C_{22}(\tau) &= 0\cdot25[\rho(\tau) + 3\rho^2(\tau)]\\
C_{23}(\tau) &= C_{32}(\tau) = \frac{0\cdot75}{12\cdot124}[-\rho(\tau) + \rho^2(\tau)]\\
C_{24}(\tau) &= C_{42}(\tau) = \frac{-1\cdot3460\rho(\tau) + 1\cdot3460\rho^2(\tau)}{49\cdot25}\\
C_{33}(\tau) &= \frac{0\cdot5625\rho(\tau) + 15\cdot1875\rho^2(\tau) + 21\rho^3(\tau)}{36\cdot75}\\
C_{34}(\tau) &= C_{43}(\tau) = \frac{1\cdot0088\rho(\tau) - 26\cdot7522\rho^2(\tau) + 25\cdot6761\rho^3(\tau)}{149\cdot2817}\\
C_{44}(\tau) &= \frac{1\cdot8097\rho(\tau) + 71\cdot1243\rho^2(\tau) + 252\cdot5364\rho^3(\tau) + 282\cdot0875\rho^4(\tau)}{606\cdot3906}
\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (49)$$

$$\left.\begin{aligned}
D_{11}(\tau) &= 0\cdot707\,1\rho(\tau)\\
D_{21}(\tau) &= 0\cdot353\,6\rho(\tau)\\
D_{31}(\tau) &= -0\cdot0875\rho(\tau)\\
D_{41}(\tau) &= -0\cdot038\,6\rho(\tau)\\
D_{51}(\tau) &= 0\cdot061\,7\rho(\tau)
\end{aligned}\right\} \quad . \quad . \quad . \quad (50)$$

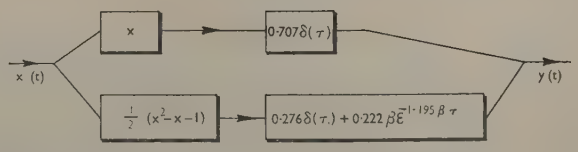


Fig. 9.—Optimum two-path memory filter for the second problem.

* The input signal, $x(t)$, has a second probability density function for which $C_{1n}(\tau) \equiv C_{n1}(\tau) \equiv 0$ [see eqn. (30)] although $D_{n1}(\tau) \neq 0$, so that the optimum filter is not linear, but the integral equation for $\omega_1(\tau)$ is independent of the rest of the integral equations.

For three parallel paths,

$$\left.\begin{aligned}\omega_1(\tau) &= 0{\cdot}7071\delta(\tau)\\ \omega(\tau) &= 0{\cdot}2766\delta(\tau) + 0{\cdot}2195\beta\varepsilon^{-1{\cdot}2128\beta\tau}\\ &\quad + 0{\cdot}0008\beta\varepsilon^{-2{\cdot}3717\beta\tau} \quad \tau \geqslant 0\\ &= 0 \quad \tau < 0\\ \omega_3(\tau) &= -0{\cdot}0624\delta(\tau) - 0{\cdot}0758\beta\varepsilon^{-1{\cdot}2128\beta\tau}\\ &\quad - 0{\cdot}0126\beta\varepsilon^{-2{\cdot}3717\beta\tau} \quad \tau \geqslant 0\\ &= 0 \quad \tau < 0\end{aligned}\right\} \quad . \quad (52)$$

The block diagrams of these filters are shown in Figs. 9 and 10.

The performance index is calculated (Fig. 11) using eqns. (105) and (113) for the zero-memory, filter and eqns. (110) and (113) for the memory filter.

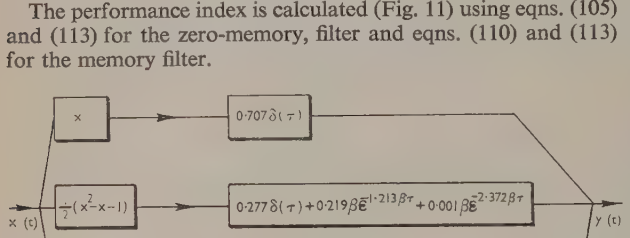


Fig. 10.—Optimum three-path memory filter for the second problem.

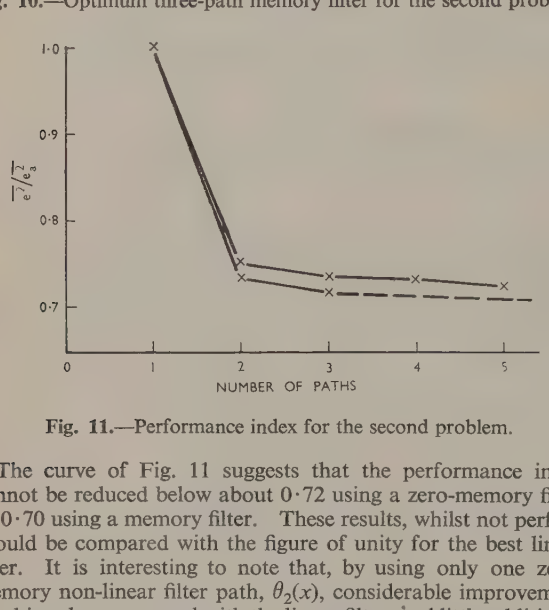


Fig. 11.—Performance index for the second problem.

The curve of Fig. 11 suggests that the performance index cannot be reduced below about 0·72 using a zero-memory filter or 0·70 using a memory filter. These results, whilst not perfect, should be compared with the figure of unity for the best linear filter. It is interesting to note that, by using only one zero-memory non-linear filter path, $\theta_2(x)$, considerable improvement is achieved as compared with the linear filter, and little additional improvement is effected by increasing the number of paths or by using a memory filter. Thus, in this example it seems that fairly rapid convergence to the optimum filter of class η_1 is obtained; this condition, whilst common, may not necessarily occur with other types of input.

(5) CONCLUSIONS

A routine procedure, unhindered by the need to evaluate difficult integrals, is proposed for the optimization of a class of memory filters which are realized as a number of parallel paths. The physical complexity of the filter increases rapidly with the number of paths, but for many applications the performance index of the filter converges rapidly to a constant value as the number of paths increases. Although the analytical optimiz tion is tedious, applications can be visualized where the labo and complexity might be justified; in particular, it has be demonstrated that, if signal and noise have the same spectr density, although a linear noise filter cannot produce a significa improvement in mean-square error, a multi-path noise filter c do so.

Multi-path zero-memory filters also can be optimized usir this procedure, which has advantages over the more dire method, since, although an optimum zero-memory filter functic can be expressed explicitly in terms of probability functio [eqn. (20)], this expression involves integrals which often ma be difficult to evaluate. In contrast to the labour involved the optimization of the multi-path memory filter, a multi-pat zero-memory filter can be optimized very easily analytical (taking about four hours for an average problem).

It is probable, in practice, that insufficient information abo the input will be available and it will be necessary to obtain th information from long samples of the input and desired outpu A method is proposed whereby the information can be obtaine in the required form by making some measurements using a analogue computer. Although no results have been obtaine in this way, no new experimental techniques are involved an the only essential requirement is a sufficiently large and accura analogue machine.

It may be necessary to employ a digital computer as well fo the solution of the set of integral equations [eqn. (26)]. Th example in the text was calculated laboriously by hand, but the seems no reason why a digital machine could not be programme to do the same task in the same way; if it should be require to solve more than two simultaneous integral equations, a digit computer would be almost indispensable.

Multi-path memory filters may be connected in cascade t produce still lower mean-square error. If an analogue com puter is available, a second multi-path memory filter can b optimized to operate on the output of an earlier filter; the mean square error generally will be reduced, because the combine filter is of higher classification[1, 2] than a filter of class η_1, an can use more statistical information about the signal and nois As the number of optimum cascaded filters is increased, th mean-square filtering error will converge to either zero or som finite value which must be less than, or equal to, the origina mean-square error at the input to the cascaded chain. Thus th last of a large number of filters in the chain will not make a appreciable difference to the mean-square error; intuitively, seems that this filter will become a direct connection, whenc in eqns. (26) and (27),

$$\left.\begin{aligned}\omega_m(\tau_2) &= 0 \quad m \neq 1\\ &= \sigma_y\delta(\tau_2) \quad m = 1\\ &= \bar{y}\delta(\tau_2) \quad m = 0\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (53$$

Therefore the coefficients of the probability-density expansio for the output and desired output satisfy

$$\sigma_y C_{n1}(\tau) \equiv D_{n1}(\tau)\sigma_z \quad . \quad . \quad . \quad . \quad . \quad (54$$
$$\bar{y} = \bar{z}$$

If the input and desired output are such that they satisfy thi equation, the optimum filter of class η_{1R} is an attenuator.

A cascade connection of two optimum zero-memory filter is itself a zero-memory filter, so that it is to be expected tha the second of the two filters will be a direct connection, sinc together they cannot produce a lower mean-square error than th first alone can achieve. However, we can cascade a zero-memor non-linear filter and a linear memory filter alternately an produce successively lower values of mean-square error. Suc

an arrangement may have considerable practical advantages owing to the relative simplicity of the optimization procedure for both zero-memory and linear memory filters.

The optimization procedure proposed in the paper can be used for the analysis of non-linear systems. Given a non-linear box of class η_1 it is always possible to construct an approximate multi-path filter of class η_{1R} (approximate only in the sense that R is finite). Experimentally, the analysis of the box can be made by measuring the cross-correlations and auto-correlation of all the outputs of a set of Hermite-polynomial generators and the non-linear box output, having connected a common Gaussian input to them all (see Fig. 12). A multi-path filter is then

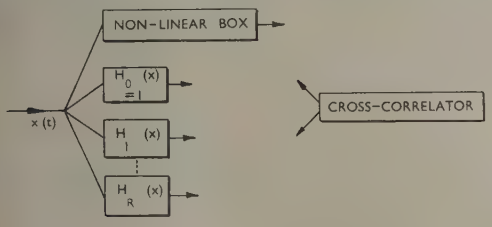


Fig. 12.—Experimental analysis of a non-linear system.

optimized to have an output as nearly as possible equal, in the mean-square-error sense, to the output of the non-linear box. The reason for the choice of Gaussian input and Hermite polynomials is that this produces integral equations [eqn. (26)] which are independent, thereby simplifying the mathematics, and also that a wide-band Gaussian input is sufficiently random for the unknown box to be completely characterized by the above measurements. Even if the unknown box is not of class η_1 an approximate equivalent filter of class η_{1R} can be found, and may be useful in the approximate analysis of more complicated non-linear systems. In this case the equivalence does not necessarily hold for different inputs and it may be better to find the equivalent filter under the same input conditions as those of the original non-linear box.

(6) ACKNOWLEDGMENTS

This work forms part of a research programme in control engineering, under the direction of Mr. J. F. Coales, Reader in Engineering at Cambridge University.

The help of many of the author's colleagues in the completion of this work through informal discussions is gratefully acknowleged; in particular, Mr. Coales has suggested many improvements both before and after reading the text; Mr. Loynes of the Statistical Laboratory suggested a generalization and extension in the theory; and a large part of the numerical calculations, which proved to be very time-consuming, was done by Mrs. Chambers and Mr. Lawrence of the Engineering Laboratory.

(7) REFERENCES

(1) ZADEH, L. A.: 'A Contribution to the Theory of Non-Linear Systems', *Journal of the Franklin Institute*, 1953, **255**, p. 387.

(2) ZADEH, L. A.: 'Optimum Non-Linear Filters', *Journal of Applied Physics*, 1953, **24**, p. 396.

(3) SINGLETON, H. A.: 'Theory of Non-Linear Transducers', M.I.T. Technical Report No. 160, 1950.

(4) BOSE, A. G.: 'A Theory of Non-Linear Systems', M.I.T. Technical Report No. 309, 1956.

(5) COURANT, R.: 'Differential and Integral Calculus', Vol. 2 (Blackie, 1936).

(6) WIENER, N.: 'Extrapolation, Interpolation and Smoothing of Stationary Time Series' (Wiley, 1949).

(7) LAMPARD, D. G.: 'A New Method of Determining Correlation Functions of Stationary Time Series', *Proceedings I.E.E.*, Monograph No. 104 R, August, 1954 (**102** C, p. 35).

(8) LUBBOCK, J. K.: 'On a Semi-Separable Class of Random Processes', *Journal of Electronics and Control* (to be published).

(9) NUTTALL, A. H.: 'Theory and Application of the Separable Class of Random Processes', M.I.T. Technical Report No. 343, 1958.

(10) BROWN, J. L.: 'On a Cross-Correlation Property for Stationary Random Processes', *Transactions of the Institute of Radio Engineers*, 1957, **IT-3**, p. 28.

(11) BARRETT, J. F., and LAMPARD, D. G.: 'An Expansion for some Second Order Probability Distributions and its Application to Noise Problems', *ibid.*, 1955, **IT-1**, p. 10.

(12) BARRETT, J. F., and COALES, J. F.: 'An Introduction to the Analysis of Non-Linear Control Systems with Random Inputs', *Proceedings I.E.E.*, Monograph No. 154 M, November, 1955 (**103** C, p. 190).

(13) BALTH, VAN DER POL, and BREMMER: 'Operational Calculus' (Cambridge, 1955).

(8) APPENDICES

(8.1) An Expansion of Second Probability Density Functions

Let $x(t)$ and $z(t)$ be random functions of time belonging to the ensembles X and Z for which the first probability density functions $\mathrm{p}(x)$ and $\mathrm{q}(z)$ are known.

Suppose there exist two sets of polynomials $\theta_n(x)$ and $\gamma_m(z)$ which are orthonormal with respect to $\mathrm{p}(x)$ and $\mathrm{q}(z)$ respectively as weighting functions, over the whole range of x and z respectively:

$$\left.\begin{aligned}\int \theta_i(x)\theta_j(x)\mathrm{p}(x)dx &= \delta_{ij}\\ \int \gamma_i(z)\gamma_j(z)\mathrm{q}(z)dz &= \delta_{ij}\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (55)$$

where

$$\begin{aligned}\delta_{ij} &= 1 \quad i = j\\ &= 0 \quad i \neq j\end{aligned}$$

The first few of these polynomials are easily identified, since

$$\left.\begin{aligned}&\int \mathrm{p}(x)dx = 1\\ &\bar{x} = \int x\mathrm{p}(x)dx\\ \text{and}\quad &\sigma_x^2 = \int (x-\bar{x})^2\mathrm{p}(x)dx\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (56)$$

where $\bar{x}$ is the mean value of $\mathrm{x}(t)$ and σ_x^2 is the variance of $\mathrm{x}(t)$. Thence we may write

$$\int 1 \times 1 \times \mathrm{p}(x)dx = 1$$

$$\int \frac{x-\bar{x}}{\sigma_x} 1 \times \mathrm{p}(x)dx = 0$$

$$\int \left(\frac{x-\bar{x}}{\sigma_x}\right)^2 \mathrm{p}(x)dx = 1$$

and we conclude that

$$\left.\begin{aligned}&\theta_0(x) = 1 \qquad \theta_1(x) = \frac{x-\bar{x}}{\sigma_x}\\ \text{and similarly}\quad &\gamma_0(z) = 1 \qquad \gamma_1(z) = \frac{z-\bar{z}}{\sigma_z}\end{aligned}\right\} \quad . \quad . \quad . \quad (57)$$

The higher-order polynomials are all obtained from the moments of the distributions $p(x)$ and $q(z)$ (see Appendix 8.2).

The joint second probability density $p(x, z; \tau)$ may be expanded in series form:

$$p_2(x, z; \tau) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} D_{nm}(\tau)\theta_n(x)\gamma_m(z)p(x)q(z) \quad . \quad (58)$$

The restrictions on $p_2(x, z; \tau)$ to validate this formal procedure will not be enumerated.

Because of the orthogonality of $\theta_n(x)$ and $\gamma_m(z)$,

$$D_{nm}(\tau) = \iint \theta_n(x)\gamma_m(z)p_2(x, z; \tau)dxdz \quad . \quad . \quad (59)$$

The coefficients D_{nm} can all be expressed in terms of the joint moments of the form

$$\alpha_{rs}(\tau) = \overline{x^r z^s} = \iint x^r z^s p_2(x, z; \tau)dxdz \quad . \quad . \quad (60)$$

and we observe from eqns. (57) and (59) that

$$D_{00} = 1 \qquad D_{0n} = D_{n0} = 0 \qquad n \neq 0 \quad . \quad . \quad (61)$$

$$D_{11} = \frac{\overline{(x - \bar{x})(z - \bar{z})}}{\sigma_x \sigma_z} = \rho_{xz}(\tau) \quad . \quad . \quad . \quad (62)$$

which is the normalized correlation function between $x(t)$ and $z(t)$. Certain other properties of these coefficients follow immediately.

$$\left.\begin{aligned} D_{nm}(\tau) &= D_{mn}(-\tau) \\ D_{nm}(0) &= \delta_{nm} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (63)$$

provided that x and z are members of the same ensemble.

(8.1.1) **Single-Series Expansion.**

There exists a class of second probability density functions,[11] Λ, of physically realizable processes for which it is possible to expand the density function in the form of a single series:

$$p_2(x, z; \tau) = \sum_{n=0}^{\infty} D_{nn}(\tau)\theta_n(x)\gamma_n(z)p(x)q(z) \quad . \quad . \quad (64)$$

An example is the Gaussian second probability density function with zero mean value, $\bar{x} = 0$, for which we may write:

$$p_2(x_1, x_2; \tau) = \frac{1}{2\pi\sqrt{(1 - \rho^2)}} \exp - \left[\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2(1 - \rho^2)}\right] . \quad (65)$$

It can be shown in this case[12] that

$$p_2(x_1, x_2; \tau) = \frac{\exp - \left(\frac{x_1^2 + x_2^2}{2}\right)}{2\pi} \sum_{n=0}^{\infty} H_n(x_1)H_n(x_2)\frac{\rho^n}{n!} \quad . \quad (66)$$

where $$H_n(x) = (-1)^n \varepsilon^{x^2/2} \frac{d^n}{dx^n} \varepsilon^{-x^2/2} \quad . \quad . \quad . \quad . \quad (67)$$

and $$\left.\begin{aligned} \int H_n(x)H_m(x)\varepsilon^{-x^2/2}dx &= n!\sqrt{(2\pi)} \quad n = m \\ &= 0 \qquad\qquad n \neq m \end{aligned}\right\} \quad . \quad . \quad (68)$$

Therefore $$\left.\begin{aligned} \theta_n(x) &= \frac{H_n(x)}{\sqrt{(n!)}} \\ p(x) &= \frac{1}{\sqrt{(2\pi)}} \varepsilon^{-x^2/2} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (69)$$

and $$D_{nn}(\tau) = \rho^n(\tau)$$

(8.2) **Analytical Determination of the Coefficients of a Second Probability Density-Function Expansion**

If both the first and second probability density functions o a process or the first and joint second probability density function of a pair of processes are known, it is usually possible to obtai both the coefficients and the polynomials of the expansion o that second probability density function or joint second proba bility density function as follows.

(8.2.1) **Moments of the Input.**

To determine the appropriate polynomials for a given proces (i.e. polynomials which are orthonormal with respect to the firs probability density function of the process as weighting function it is necessary first to determine the moments of that process These can be found from the characteristic function $\psi(u)$ of th random process, which is defined as

$$\psi(u) = E(\varepsilon^{jux}) = \int \varepsilon^{jux} p(x)dx \quad . \quad . \quad . \quad (70$$

Thus $\psi(u)$ is the Fourier transform of $p(x)$.

From the characteristic function the moments of the proces may be determined using the following formula for the nt moment:

$$\alpha_n = \overline{x^n} = (-1)^{-n/2}\left[\frac{d^n}{du^n}\psi(u)\right]_{u=0} \quad . \quad . \quad . \quad (71$$

(*a*) In filtering problems the filter input often consists of th sum of a signal $z(t)$ and noise $v(t)$:

$$x(t) = z(t) + v(t) \quad . \quad . \quad . \quad . \quad . \quad . \quad (72$$

whence $$\alpha_n = \overline{x^n} = \overline{(v + z)^n}$$

$$= \sum_{r=0}^{n} {}^nC_r \overline{v^{n-r}}\,\overline{z^r} \quad . \quad . \quad . \quad . \quad . \quad (73$$

when $z(t)$ and $v(t)$ are statistically independent.

(*b*) The input may be equal to the product of signal and nois

$$x(t) = z(t)v(t) \quad . \quad . \quad . \quad . \quad . \quad . \quad (74$$

In this case the moments of the input are found more easil since

$$\overline{x^n} = \overline{z^n v^n} = \overline{z^n}\,\overline{v^n} \quad . \quad . \quad . \quad . \quad . \quad (75$$

when $z(t)$ and $v(t)$ are statistically independent.

(8.2.2) **Orthonormal Polynomials.**

An orthonormal polynomial of order n will consist of a linea summation of the first n of the linearly independent functio

$$1, x, x^2, x^3, \ldots$$

Putting $\theta_0(x) = 1$ and $\theta_1(x) = a_{11}x - a_{10}$ such that

$$\int \theta_0(x)\theta_1(x)p(x)dx = \int (a_{11}x - a_{10})p(x)dx = 0$$

and $$\int \theta_1^2(x)p(x)dx = \int (a_{11}x - a_{10})^2 p(x)dx = 1$$

we find that $a_{11} = 1/\sigma_x$ and $a_{10} = \bar{x}/\sigma_x$.

Therefore $$\theta_1(x) = \frac{x - \bar{x}}{\sigma_x} \quad . \quad . \quad . \quad . \quad . \quad (76$$

Let $$\theta_2(x) = a_{22}x^2 - a_{21}\theta_1(x) - a_{20}$$

such that

$$\int \theta_2(x)\theta_1(x)\mathrm{p}(x)dx = \int [a_{22}x^2 - a_{21}\theta_1(x) - a_{20}]\theta_1(x)\mathrm{p}(x)dx = 0$$

$$\int \theta_2(x)\theta_0(x)\mathrm{p}(x)dx = \int [a_{22}x^2 - a_{21}\theta_1(x) - a_{20}]\mathrm{p}(x)dx = 0$$

and $$\int \theta_2^2(x)\mathrm{p}(x)dx = \int [a_{22}x^2 - a_{21}\theta_1(x) - a_{20}]^2\mathrm{p}(x)dx = 1$$

Then
$$\theta_2(x) = \frac{x^2 - \overline{x^2\theta_1(x)}\theta_1(x) - \overline{x^2}}{\left\{\overline{x^4} - [\overline{x^2\theta_1(x)}]^2 - [\overline{x^2}]^2\right\}^{1/2}} \quad . \quad . \quad (77)$$

For the general case

$$\theta_n(x) = a_{nn}x^n - \sum_{r=0}^{n-1} a_{nr}\theta_r(x) \quad . \quad . \quad . \quad . \quad (78)$$

where
$$\left.\begin{aligned} a_{nr} &= a_{nn}\overline{x^n\theta_r(x)} \\ a_{nn} &= \left\{\overline{x^{2n}} - \sum_{r=0}^{n-1}[\overline{x^n\theta_r(x)}]^2\right\}^{-1/2} \end{aligned}\right\} \quad . \quad . \quad . \quad (79)$$

Therefore
$$\theta_n(x) = \frac{x^n - \sum_{r=0}^{n-1}\overline{x^n\theta_r(x)}\theta_r(x)}{\left\{\overline{x^{2n}} - \sum_{r=0}^{n-1}[\overline{x^n\theta_r(x)}]^2\right\}^{1/2}} \quad . \quad . \quad . \quad (80)$$

(8.2.3) Coefficients of the Second Probability Density-Function Expansions.

Using eqns. (79) we can obtain a recurrence formula for the coefficients $D_{n1}(\tau)$ occurring in the expansion of the joint second probability density $\mathrm{p}_2(x, z; \tau)$ as follows:

$$\sigma_z D_{n1}(\tau) = \sigma_z\overline{\theta_n(x)\gamma_1(z)} = \overline{\theta_n(x)z} \qquad n \neq 0$$
$$= 0 \qquad n = 0$$

Therefore
$$\sigma_z D_{n1}(\tau) = a_{nn}\overline{x^n z} - \sum_{r=1}^{n-1} a_{nr}\theta_r(x)z$$

i.e.
$$\sigma_z D_{n1}(\tau) = a_{nn}\lambda_{n1} - \sum_{r=1}^{n-1} a_{nr}\sigma_z D_{r1}(\tau) \quad . \quad . \quad (81)$$

where
$$\lambda_{n1}(\tau) = \iint x^n z\mathrm{p}_{xz}(x, z; \tau)dxdz \quad . \quad . \quad . \quad (82)$$

A similar but more complicated expression can be found for the coefficients $C_{nm}(\tau)$ occurring in the expansion of the second probability density:

$$\mathrm{p}_2(x_1, x_2; \tau) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} C_{nm}(\tau)\theta_n(x_1)\theta_m(x_2)\mathrm{p}(x_1)\mathrm{p}(x_2) \quad . \quad (83)$$

where
$$C_{nm}(\tau) = \iint \theta_n(x_1)\theta_m(x_2)\mathrm{p}_2(x_1, x_2; \tau)dx_1dx_2 \quad . \quad (84)$$

For convenience we choose to express eqn. (80) in the form

$$\theta_n(x) = b_n x^n + b_{n-1}x^{n-1} + \ldots + b_r x^r + \ldots + b_0 \quad . \quad (85)$$

Then
$$C_{nm}(\tau) = \sum_{r=0}^{n}\sum_{s=0}^{m} b_r b_s \alpha_{rs}(\tau) \quad . \quad . \quad . \quad . \quad (86)$$

where
$$\alpha_{rs}(\tau) = \iint x_1^r x_2^s \mathrm{p}_{xx}(x_1, x_2; \tau)dx_1dx_2 \quad . \quad . \quad (87)$$

Therefore all the coefficients $C_{nm}(\tau)$ and $D_{n1}(\tau)$ needed for the optimization can be expressed in terms of the joint moments $\alpha_{rs}(\tau)$ and $\lambda_{n1}(\tau)$ defined by eqns. (87) and (82) respectively. These joint moments can be found from the joint characteristic functions; in particular

$$\lambda_{nm}(\tau) = \overline{x^n z^m} = (-1)^{-\frac{(n+m)}{2}}\frac{\partial^n}{\partial u_1^n}\frac{\partial^m}{\partial u_2^m}\psi_{xz}(u_1, u_2; \tau) \quad . \quad (88)$$

where
$$\psi_{xz}(u_1, u_2; \tau) = \iint \varepsilon^{j(u_1x+u_2z)}\mathrm{p}_{xz}(x, z; \tau)dxdz \quad . \quad (89)$$

(*a*) When the input is equal to the sum of signal, $z(t)$ and noise, $v(t)$,

$$\alpha_{nm} = \overline{x_1^n x_2^m} = \overline{(v_1 + z_1)^n(v_2 + z_2)^m}$$
$$= \sum_{r=0}^{n}\sum_{s=0}^{m} {}^nC_r\,{}^mC_s\overline{v_1^r z_1^{n-r}v_2^s z_2^{m-s}}$$

and when $z(t)$ and $v(t)$ are statistically independent,

$$\alpha_{nm} = \sum_{r=0}^{n}\sum_{s=0}^{m} {}^nC_r\,{}^mC_s\overline{z_1^{n-r}z_2^{m-s}}\,\overline{v_1^r v_2^s} \quad . \quad . \quad . \quad (90)$$

$$\lambda_{n1} = \overline{x^n z} = \overline{(v_1 + z_1)^n z_2}$$
$$= \sum_{r=0}^{n} {}^nC_r\overline{v_1^r z_1^{n-r}z_2}$$

and when $z(t)$ and $v(t)$ are statistically independent,

$$\lambda_{n1} = \sum_{r=0}^{n} {}^nC_r\overline{v_1^r}\;\overline{z_1^{n-r}z_2} \quad . \quad . \quad . \quad . \quad (91)$$

(*b*) When the input is equal to the product of signal and noise the joint moments are easily found, since

$$\alpha_{nm}(\tau) = \overline{x_1^n x_2^m} = \overline{z_1^n v_1^n z_2^m v_2^m}$$
$$= \overline{z_1^n z_2^m}\,\overline{v_1^n v_2^m} \quad . \quad . \quad . \quad . \quad . \quad . \quad (92)$$

provided that $z(t)$ and $v(t)$ are statistically independent, and

$$\lambda_{n1}(\tau) = \overline{x_1^n z_2} = \overline{z_1^n z_2}\;\overline{v_1^n} \quad . \quad . \quad . \quad . \quad (93)$$

provided that $z(t)$ and $v(t)$ are statistically independent.

It will be appreciated that the labour involved may be very great in order to compute first the polynomials and then the coefficients of an expansion for the second or joint probability density function from a knowledge of either the second or joint characteristic function or the probability density function itself. Also, when the expansion corresponds to a process which itself is the sum of two other processes whose probability densities are given [as in (*a*)] the work is greater still. One such problem, the solution to which is given earlier, took about 120 hours to calculate the expansion coefficients on a desk calculator.

(8.3) Mean-Square Filtering Error

The mean-square filtering error is given by eqn. (1):

$$\overline{e^2} = \overline{[y(t) - z(t)]^2} = \overline{y(t)^2} + \overline{z(t)^2} - 2\overline{y(t)z(t)} \quad . \quad (1)$$

Let $y(t) = \mathrm{g}[x(t)]$, where g represents any linear or non-linear operation, of finite mean square, on the whole, or part, of the past input.

The variation in $\overline{e^2}$ produced by a variation in $\mathrm{g}[x]$ of $\epsilon\mathrm{h}[x]$, where ϵ is a small constant and $\mathrm{h}[x]$ again depends only on the past of x, is given by

$$\Delta\overline{e^2} = -2\epsilon\overline{\mathrm{h}(x)[\mathrm{g}(x) - z]} + \epsilon^2\overline{\mathrm{h}(x)^2}$$

For g(x) to be optimum,

$$\left[\frac{\partial \overline{\Delta e^2}}{\partial \epsilon}\right]_{\epsilon=0} = -2\overline{h(x)[g(x) - z]} = 0$$

Therefore $$\overline{h(x)[g(x) - z]} = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (94)$$

Thence the following two conditions are derived:

For a chosen variation $\epsilon h(x) = \epsilon g(x)$

$$\overline{g(x)^2} = \overline{zg(x)}$$

i.e. $$\overline{y(t)^2} = \overline{z(t)y(t)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (95)$$

For a chosen variation $\epsilon h(x) = \epsilon \times 1$

$$\overline{g(x)} = \bar{z}$$

i.e. $$\overline{y(t)} = \overline{z(t)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (96)$$

Eqns. (95) and (96) are seen to be necessary if g(x) represents an optimum filter belonging to any class.

Substitution of eqns. (95) and (96) in eqn. (1) yields

$$\overline{e^2} = \sigma_z^2 - \sigma_y^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (97)$$

That this result is true for each of the classes of filters previously considered is easily verified.

Furthermore, suppose there should exist more than one optimum solution for g(x); by taking any two solutions $g_1(x)$ and $g_2(x)$ we may write, using eqn. (94),

$$\overline{g_1(x)[g_2(x) - z]} = 0$$

Similarly $$\overline{g_2(x)[g_1(x) - z]} = 0$$

whence $$\overline{g_1(x)z} = \overline{g_2(x)z} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (98)$$

From eqns. (95) and (98) it is seen that the mean-square error [eqn. (1)] is the same for each of the two solutions $g_1(x)$ and $g_2(x)$. Thus although a particular solution may not be unique we can be sure that it is no worse than any other solution in so far as it gives the same mean-square error.

(8.3.1) The Class of Multi-Path Zero-Memory Filters of Degree *R*.

Eqn. (94) gives the sufficient conditions for g(x) to represent an optimum filter of any specified class where h(x) represents any physically realizable filter of the same class.

For the class of multi-path zero-memory filters of degree R,

$$g(x) = \sum_{m=0}^{R} k_{gm}\theta_m(x)$$

The most general form for h(x) is

$$h(x) = \sum_{n=0}^{R} k_{hn}\theta_n(x)$$

Substitution in eqn. (94) gives

$$\sum_{n=0}^{R} k_{hn}\left[\sum_{m=0}^{R} k_{gm}\overline{\theta_n(x)\theta_m(x)} - \overline{\theta_n(x)z}\right] = 0 \quad . \quad . \quad (99)$$

Since $$\overline{\theta_n(x)\theta_m(x)} = 0 \quad m \neq n$$
$$= 1 \quad m = n$$

we obtain $$\sum_{n=0}^{R} k_{hn}k_{gn} = \sum_{n=0}^{R} k_{hn}\overline{\theta_n(x)z} \quad . \quad . \quad . \quad . \quad (100)$$

Since the set of constants k_{hn} can be given any pre-assigned se of values only one solution exists, namely

$$k_{gn} = \overline{\theta_n(x)z} \quad . \quad . \quad . \quad . \quad . \quad . \quad (102$$
$$n = 0, 1, 2 \ldots R$$

Also, since $\theta_0(x) = 1$,

$$k_{g0} = \overline{g(x)} = \overline{y(t)} = \overline{z(t)} \quad . \quad . \quad . \quad . \quad (103$$

and the optimum multi-path zero-memory filter of degree R i given by

$$f[x(t)] = \sum_{n=1}^{R} \sigma_z D_{n1}\theta_n[x(t)] + \overline{z(t)} \quad . \quad . \quad . \quad (104$$

This equation is identical with the equation obtained by trun cating the expansion (21). From eqn. (104)

$$\sigma_y^2 = \overline{\{f[x(t)] - \overline{z(t)}\}^2} = \sum_{n=1}^{R} \sigma_z^2 D_{n1}^2$$

since $$\overline{\theta_n(x)\theta_m(x)} = 0 \quad m \neq n$$
$$= 1 \quad m = n$$

Substitution in eqn. (97) gives

$$\overline{e^2} = \sigma_z^2\left(1 - \frac{\sigma_y^2}{\sigma_z^2}\right) = \sigma_z^2\left(1 - \sum_{n=1}^{R} D_{n1}^2\right) \quad . \quad . \quad (105$$

(8.3.2) The Class of Multi-Path Memory Filters of Degree *R*.

$$g(x) = \sum_{m=0}^{R}\int_0^\infty \omega_{gm}(\tau)\theta_m[x(t - \tau)]d\tau$$

The most general form for h(x) is

$$h(x) = \sum_{n=0}^{R}\int_0^\infty \omega_{hn}(\tau)\theta_n[x(t - \tau)]d\tau$$

Substitution in eqn. (94) gives

$$\sum_{n=0}^{R}\int_0^\infty \omega_{hn}(\tau_1)\left\{\sum_{m=0}^{R}\int_0^\infty \omega_{gm}(\tau_2)C_{nm}[(\tau_1 - \tau_2)]d\tau_2 - \overline{\theta_n[x(t - \tau_1)]z(t)}\right\}d\tau_1 = 0 \quad . \quad . \quad (106$$

Using the fundamental lemma of variation calculus[5] and notin that all $\omega_{hn}(\tau)$ are arbitrary but physically realizable weightin functions of linear networks [i.e. $\omega_{hn}(\tau) = 0$ for $\tau < 0$], w obtain

$$\sum_{m=0}^{R}\int_0^\infty \omega_m(\tau_2)C_{nm}[(\tau_1 - \tau_2)]d\tau_2 - \overline{\theta_n[x(t - \tau_1)z](t)} = 0 \quad (107$$
$$n = 0, 1, 2 \ldots R \quad \tau_1 \geqslant 0$$

but since $C_{0m} = 0$, $m \neq 0$; $C_{0m} = 1$, $m = 0$; and $\theta_0[x(t)] = 1$

$$\int_0^\infty \omega_0(\tau)d\tau = \omega_0 = \overline{z(t)} \quad . \quad . \quad . \quad . \quad (27$$

and $$\sum_{m=1}^{R}\int_0^\infty \omega_m(\tau_2)C_{nm}[(\tau_1 - \tau_2)]d\tau_2 = \sigma_z \quad D_{n1}(\tau_1) \quad \tau_1 \geqslant 0$$
$$. \quad . \quad . \quad (26$$

From eqns. (26) and (27), it follows that

$$\sigma_y^2 = \sum_{n=1}^{R}\sum_{m=1}^{R}\int_0^\infty\int_0^\infty \omega_n(\tau_1)\omega_m(\tau_2)C_{nm}[(\tau_1 - \tau_2)]d\tau_1 d\tau_2 \quad (108$$

or we may write $$\sigma_y^2 = \sum_{n=1}^{R}\int_0^\infty D_{n1}(\tau)\sigma_z\omega_n(\tau)d\tau \quad . \quad . \quad . \quad (109$$

and by substitution in eqn. (97) obtain:

$$\overline{e^2} = \sigma_z^2\left(1 - \frac{\sigma_y^2}{\sigma_z^2}\right) = \sigma_z^2 - \sum_{n=1}^{R}\int_0^\infty D_{n1}(\tau)\sigma_z\omega_n(\tau)d\tau \quad . \quad (110)$$

(8.3.3) Performance Index for Noise Filters.

Even in the most unfavourable conditions when the statistical characteristics of the signal and noise are identical, an optimum linear or non-linear noise filter will reduce the mean-square output error given by the optimum attenuator; for example, the optimum linear filter for the condition when the input signal and noise have the same spectral densities becomes an attenuator. Therefore it is both logical and convenient to choose the mean-square error of the optimum attenuator as a datum against which to compare the mean-square errors of more complex optimum filters.

The optimum attenuator mean-square error is given by eqn. (107):

$$\overline{e_a^2} = \sigma_z^2(1 - D_{11}^2) \quad . \quad . \quad . \quad . \quad . \quad (111)$$

and the comparative mean-square error by

$$\frac{\overline{e^2}}{\overline{e_a^2}} = \frac{1 - \sigma_y^2/\sigma_z^2}{1 - D_{11}^2} \quad . \quad . \quad . \quad . \quad . \quad (112)$$

When the signal and noise are additive and uncorrelated,

$$\frac{\overline{e^2}}{\overline{e_a^2}} = \frac{1 - \sigma_y^2/\sigma_z^2}{1 - \sigma_z^2/\sigma_x^2} \quad . \quad . \quad . \quad . \quad . \quad (113)$$

where $\overline{e_a^2}$ is the mean-square error of the optimum attenuator, $\overline{e^2}$ is the mean-square filter error and $0 \leqslant \overline{e^2}/\overline{e_a^2} \leqslant 1$.

The two limiting values of $\overline{e^2}/\overline{e_a^2}$ represent the extreme conditions when the optimum filter is perfect and when it is no better than the optimum attenuator. If $\overline{e^2}/\overline{e_a^2} < 1$ we say the improvement in mean-square error is 'significant'.

(8.4) Method of Undetermined Coefficients

The solution for the set of integral equations [eqn. (26)] has been worked out by Wiener,[4] and a modification of his method is formally presented here.

The integral equations are

$$\sum_{m=1}^{R}\int_0^\infty C_{nm}[(\tau_1 - \tau_2)]\omega_m(\tau_2)d\tau_2 - D_{n1}(\tau_1)\sigma_z = H_n(\tau_1) \quad (26)$$

where $H_n(\tau_1) = 0 \quad \tau_1 \geqslant 0 \quad n = 1, 2 \ldots R$

Taking the bilateral Laplace transform[13] of both sides of these equations,

$$\sum_{m=1}^{R}\mathscr{C}_{nm}(p)\mathscr{W}_m(p) - \mathscr{D}_{n1}(p) = \mathscr{H}_n(p) \quad . \quad . \quad (114)$$

where $\mathscr{H}_n(p)$ has singularities in the right-hand half of the p-plane only [since $H_n(\tau_1) = 0$, $\tau_1 \geqslant 0$].

For simplicity the method will be described for the case of two equations only.

Let
$$\Phi(p) = \begin{vmatrix} \mathscr{C}_{11}(p) & \mathscr{C}_{12}(p) \\ \mathscr{C}_{21}(p) & \mathscr{C}_{22}(p) \end{vmatrix} \quad . \quad . \quad . \quad . \quad (115)$$

then
$$\Phi(p)\mathscr{W}_1(p) = \begin{vmatrix} \mathscr{D}_{11}(p) + \mathscr{H}_1(p) & \mathscr{C}_{12}(p) \\ \mathscr{D}_{21}(p) + \mathscr{H}_2(p) & \mathscr{C}_{22}(p) \end{vmatrix} \quad . \quad (116)$$

and
$$\Phi(p)\mathscr{W}_2(p) = \begin{vmatrix} \mathscr{C}_{11}(p) & \mathscr{D}_{11}(p) + \mathscr{H}_1(p) \\ \mathscr{C}_{21}(p) & \mathscr{D}_{21}(p) + \mathscr{H}_2(p) \end{vmatrix} \quad . \quad (117)$$

Therefore

$$\Phi(p)\mathscr{W}_1(p) = [\mathscr{D}_{11}\mathscr{C}_{22} - \mathscr{D}_{21}\mathscr{C}_{12}] + \mathscr{H}_1\mathscr{C}_{22} - \mathscr{H}_2\mathscr{C}_{12} \quad . \quad (118)$$

Suppose we can write $\Phi(p) = \Phi^+(p)\Phi^-(p)$ . . . . (119)

where $\Phi^+(p)$ has poles and zeros in the left-hand half of the p-plane only, and $\Phi^-(p)$ has poles and zeros in the right-hand half only. Then

$$\mathscr{W}_1(p)\Phi^+(p) = \frac{\mathscr{D}_{11}\mathscr{C}_{22} - \mathscr{D}_{21}\mathscr{C}_{12}}{\Phi^-(p)} + \frac{\mathscr{H}_1\mathscr{C}_{22} - \mathscr{H}_2\mathscr{C}_{12}}{\Phi^-(p)} \quad . \quad (120)$$

$\mathscr{W}_1(p)\Phi^+(p)$ has poles in the left-hand half of the p-plane only if $\mathscr{W}_1(p)$ is the transfer function of a stable system, i.e. $\omega_1(\tau) = 0, \tau < 0$. Also, it may usually be assumed that

$$\lim_{p\to\infty}\mathscr{W}_1(p)\Phi^+(p) = 0 \quad . \quad . \quad . \quad . \quad (121)$$

It may be necessary to impose some further restriction on $\mathscr{W}_m(p)$ so as to obtain a unique solution to eqns. (26). This restriction is of the form:

$$\lim_{p\to\infty}\frac{\mathscr{W}_m(p)}{p^i} = \text{constant} \quad . \quad . \quad . \quad . \quad (122)$$

where for noise-filtering problems normally $i = 0$ and in prediction problems normally $i = 1$. Only approximate physical filters for $i \geqslant 0$ can be realized, but the approximation is not serious in practice if i is not too large. In any particular problem i is chosen so that a unique solution is obtained.

Assuming $\mathscr{W}_1(p)\Phi^+(p)$ is a rational algebraic function, its numerator must of a lower order than its denominator.

Therefore

$$\mathscr{W}_1(p)\Phi^+(p) = \left[\frac{\mathscr{D}_{11}\mathscr{C}_{22} - \mathscr{D}_{21}\mathscr{C}_{12}}{\Phi^-(p)}\right]_+ + \left[\frac{\mathscr{H}_1\mathscr{C}_{22} - \mathscr{H}_2\mathscr{C}_{12}}{\Phi^-(p)}\right]_+ \quad . \quad . \quad . \quad . \quad (123)$$

where the suffixes '+' indicate that only those terms of the partial-fraction expansion representing poles in the left-hand half of the p-plane have been taken.

Since $\mathscr{H}_1$ and $\mathscr{H}_2$ have no poles in the left-hand half-plane, poles in this half-plane of $(\mathscr{H}_1\mathscr{C}_{22} - \mathscr{H}_2\mathscr{C}_{12})/\Phi^-(p)$ must be poles of both $\mathscr{C}_{22}$ and $\mathscr{C}_{12}$. We can write

$$\frac{\mathscr{H}_1\mathscr{C}_{22} - \mathscr{H}_2\mathscr{C}_{12}}{\Phi^-(p)} = \frac{a}{p+\alpha} + \frac{b}{p+\beta} + \frac{c}{p+\gamma} \quad . \quad . \quad (124)$$

where $a, b, c, \ldots$ are undetermined coefficients, which may be determined by substitution for $\mathscr{W}_1(p)$ and $\mathscr{W}_2(p)$ in one of the original equations (114).

A set of simultaneous equations in the undetermined coefficients of both $\mathscr{W}_1$ and $\mathscr{W}_2$ is obtained by equating residues of left-hand half-plane poles of this equation and by using eqn. (122).

(8.5) A Second Probability Density of Class Λ

We shall consider the signal developed at the output of a zero-memory non-linear device characterized by the equation

$$z(t) = \frac{\mathrm{H}_2[s(t)]}{\sqrt{2}} = \frac{s^2 - 1}{\sqrt{2}} \quad . \quad . \quad . \quad . \quad (125)$$

where $z(t)$ is the output, $s(t)$ the input

and
$$\mathrm{H}_n(s) = (-1)^n\varepsilon^{s^2/2}\frac{d^n}{ds^n}\varepsilon^{-s^2/2}$$

When the input is Gaussian with zero mean and unit variance, the output is a signal whose probability density function $p_2(z_1, z_2; \tau)$ is of class Λ.

It is always possible to find a polynomial θ_m of order m such that

$$y = \theta_m[H_2(s)] = H_{2m}(s) \quad . \quad . \quad . \quad . \quad (126)$$

A proof of this statement is not given here but it can be verified for all finite values of m as follows:

Let θ_m and θ_n be two such polynomials, and write

$$\phi_{mn} = \iint \theta_m(z_1)\theta_n(z_2)p_2(z_1, z_2; \tau)dz_1dz_2$$

where ϕ_{mn} is the cross-correlation between the outputs y_1 and y_2 of the polynomial-function generators θ_n and θ_m (see Fig. 13).

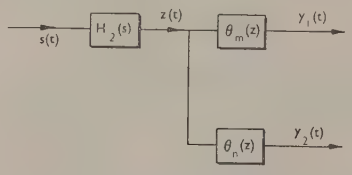


Fig. 13.—Polynomials θ_n and θ_m chosen so that $y_1 = H_{2n}[s(t)]$ and $y_2 = H_{2m}[s(t)]$ when $s(t)$ is a Gaussian signal.

Then

$$\phi_{mn} = \iint \theta_m[H_2(s_1)]\theta_n[H_2(s_2)]p_2(s_1, s_2; \tau)ds_1ds_2$$

$$= \iint H_{2m}(s_1)H_{2n}(s_2)p_2(s_1, s_2; \tau)ds_1ds_2$$

whence

$$\begin{aligned} \phi_{mn} &= 0 \qquad n \neq m \\ &= \rho_{s_1s_2}^{2n}(\tau) \qquad n = m \end{aligned} \quad . \quad . \quad . \quad (127)$$

where $\rho_{s_1s_2}$ is the normalized auto-correlation coefficient of $s(t)$. But $\phi_{mn} = C_{mn}$, using the same notation as used in eqns. (86) and (87), and $p_2(z_1, z_2; \tau)$ is a member of class Λ. We also note that $\rho_{z_1z_2}(\tau) = \rho_{s_1s_2}^2(\tau)$.

Therefore

$$C_{nn}(\tau) = \rho_{z_1z_2}^n(\tau) \quad . \quad . \quad . \quad . \quad . \quad (128)$$

whence, if $\rho_{s_1s_2}(\tau) = \varepsilon^{-\beta|\tau|}$, then $\rho_{z_1z_2}(\tau) = \varepsilon^{-2\beta|\tau|}$. The probability density of z can be obtained as follows:

$$p(s) = \frac{1}{2\pi}\varepsilon^{-s^2/2} \text{ for zero mean and unit variance.}$$

$$s = \sqrt{(z\sqrt{2} + 1)}$$

$$2sds = 2dz$$

Therefore

$$\begin{aligned} p(z) &= \frac{1}{2\sqrt{\pi}}\varepsilon^{-\sqrt{2}+1/2}\frac{1}{\sqrt{(z\sqrt{2}+1)}} \qquad z > -\frac{1}{\sqrt{2}} \\ &= 0 \qquad z \leqslant -\frac{1}{\sqrt{2}} \end{aligned} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (129)$$

The moments of z can be found from

$$\overline{z^n} = \int z^n p(z)dz$$

or from

$$\overline{z^n} = \left(\frac{\overline{s^2 - 1}}{\sqrt{2}}\right)^n \quad . \quad . \quad . \quad . \quad . \quad (130)$$

given the moments of a Gaussian process with zero mean and unit variance.

Whence $\bar{z} = 0$, $\overline{z^2} = 1$, $\overline{z^3} = 2{\cdot}828\,427$, $\overline{z^4} = 15$, $\overline{z^5} = 96{\cdot}166\,522$, $\overline{z^6} = 755{\cdot}0$, $\overline{z^7} = 6\,983{\cdot}386\,571$, $\overline{z^8} = 74\,417$, $\overline{z^9} = 897\,799{\cdot}337\,9$ and $\overline{z^{10}} = 12\,096\,873$.

621.316.57: 621.316.923

The Institution of Electrical Engineers
Monograph No. 345 U
Oct. 1959
©

DISCRIMINATION BETWEEN H.R.C. FUSES AND MINIATURE CIRCUIT-BREAKERS

By H. D. EINHORN, Dr.-Ing., Ph.D., Associate Member.

(*The paper was first received 9th March, and in revised form 8th June,* 1959. *It was published as an* INSTITUTION MONOGRAPH *in October,* 1959.)

SUMMARY

After a review of the aims and merits of protective devices for low-voltage installations, the joint use of h.r.c. fuses and miniature circuit-breakers is shown to be functional in many installations. The hypothesis that the 0·01 sec fusing current forms a criterion for discrimination is subjected to experimental tests and the statistical nature of discrimination is analysed. Practical conclusions for installation design are drawn, and a new fuse series and various other methods for backing-up miniature circuit-breakers are suggested.

(1) AIMS AND MEANS OF PROTECTION

Two trends have developed in the protection of low-voltage installations in recent years: the h.r.c. fuse has found wide application in England, while the miniature circuit-breaker has become popular in America. Both these trends have reached South Africa, where, at present, h.r.c. fuses and miniature circuit-breakers compete in replacing the semi-enclosed fuse in better-class installations. This has raised problems of practical layout design which led to the investigations presented in this and a previous paper.[1]

In order to assess the merits of different devices, a brief glance at the requirements of protective systems is warranted; these are

(*a*) Short-circuits must be cleared rapidly by protection devices of adequate breaking capacity, and overloads must be disconnected after a suitable time.

(*b*) When faults occur, interruption of supply should be localized by suitable discrimination, and reinstatement should be rapid and convenient.

Since circuits are numerous in l.v. installations and less important than the main arteries of a big distribution network, l.v. protective devices should be relatively cheap; discrimination is achieved by fuses or circuit-breakers with simple delay mechanisms, usually of an inverse time characteristic type, namely*

(*a*) *The high rupturing capacity fuse* (h.r.c. fuse)—a cartridge fuse with a breaking capacity of at least 15 kA (B.S. class AC 3, 4, or 5).

(*b*) *The semi-enclosed fuse*—a rewirable fuse with a breaking capacity below 4 kA.

(*c*) *The miniature circuit-breaker* (m.c.b.)—a small air circuit-breaker with a breaking capacity not exceeding 5 kA.

(*d*) *The air circuit-breaker*, with a breaking capacity of at least 10 kA.

(*e*) *The oil circuit-breaker*, with a breaking capacity of at least 10 kA.

All these devices can claim a place in different installations and, for some applications, several can compete at present. It is the aim of the paper to discuss rational principles of selection, to present recent tests on the discrimination of m.c.b.'s and h.r.c. fuses, and to establish a theoretical basis for interpreting the results and predicting discrimination.

* These definitions are convenient for the purposes of the paper, although the term 'm.c.b.' may be used elsewhere for moulded air circuit-breakers of 10 kA breaking capacity and although oil circuit-breakers of less than 10 kA breaking capacity are being made.
'L.V. installation' here defines a 3-phase 400-volt or single-phase 230-volt installation. 'Rated breaking capacities' are given in r.m.s. amperes at these voltages, while 'slow' and 'fast' h.r.c. fuses are defined in Section 10.1.

Correspondence on Monographs is invited for consideration with a view to publication.
Dr. Einhorn is Senior Lecturer in Electrical Engineering, University of Cape Town.

(2) SELECTION OF PROTECTIVE DEVICES

(2.1) Merits of Various Protective Devices

Circuit-breakers are fundamentally more convenient than fuses and allow easy and rapid reinstatement of circuits after a fault has been cleared. They are popular for this reason, particularly in residential and commercial premises. Protection against sustained small overloads by circuit-breakers which open at 1·25 times rated load compares favourably with the fusing factors commonly attainable.

The breaking capacities of m.c.b.'s, however, are very much lower than those of h.r.c. fuses, with which they compete on a basis of first cost, while both oil and air circuit-breakers are relatively expensive.

H.R.C. fuses have excellent breaking capacities and compare favourably with semi-enclosed fuses as regards precision and fusing factors (about 1·4–1·6).

Semi-enclosed fuses commend themselves by cheapness, and discriminate easily in front of m.c.b.'s, but they have a low breaking capacity and are liable to deteriorate, which necessitates a fusing factor of 2. They also lend themselves to easy abuse.

(2.2) Selection for Breaking Capacity

The most critical requirement of a sound installation is adequate breaking capacity. But breaking capacities greatly in excess of possible fault currents seem to offer little benefit, and convenience or cost become decisive factors where short-circuits are limited. A rational choice of equipment will therefore depend greatly on its position and on the duty it is expected to perform. A previous analysis of short-circuit currents in typical l.v. installations[1] can be summarized as follows:

(*a*) Short-circuits at or near the distribution boards of a large installation usually exceed the breaking capacity of m.c.b.'s or semi-enclosed fuses, but are rare.

(*b*) The majority of faults occur at the load end. They are appreciably limited by the resistance of the load-end cables from sub-distribution board to final load-outlets (LC in Fig. 1).

(*c*) The highest short-circuit current due to a load-end fault depends very much on the length of this final load cable. Fig. 2 illustrates the following:

(i) Almost all single-phase lighting circuits wired with 3/·036 in cables or less are within the capacity of a 5 kA m.c.b., and wiring distances in the order of 10 ft are sufficient to bring the fault current below 2·5 kA.

(ii) Most single-phase power circuits wired with 7/·036 cables can also be controlled by a 5 kA m.c.b., but minimum distances in the order of 20–30 ft are required for a 2·5 kA m.c.b. (Earth faults in 3-phase circuits are not likely to be more severe than faults in single-phase circuits.)

(iii) Phase-to-phase or 3-phase faults, however, can be much heavier, and the minimum distance even for a 5 kA m.c.b. would be 20–30 ft.

(*d*) The short-circuit phase angle of most load-end faults is small, since they are mainly limited by cable resistance; tests have

shown that the breaking capacity of a circuit-breaker is in this case relatively high.[3]

(*e*) Behind a small transformer (of 200 kVA or less), even a fault at the main distribution board will not exceed 5 kA; the phase angle of a fault which is mainly limited by transformer impedance is, however, large.

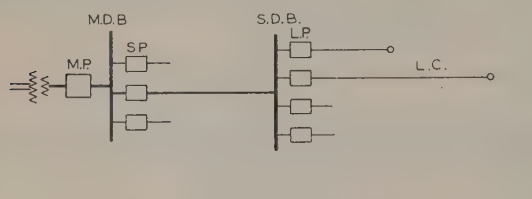


Fig. 1.—Circuit of typical l.v. installation.

M.D.B. = Main distribution board.
S.D.B. = Sub-distribution board.
M.P. = Main protection.
S.P. = Sub-mains protection.
L.P. = Load circuit protection.
L.C. = Load circuit.

Lower portion shows protection ranges
(i) with m.c.b.'s on sub-distribution boards.
(ii) with h.r.c. fuses on sub-distribution boards.

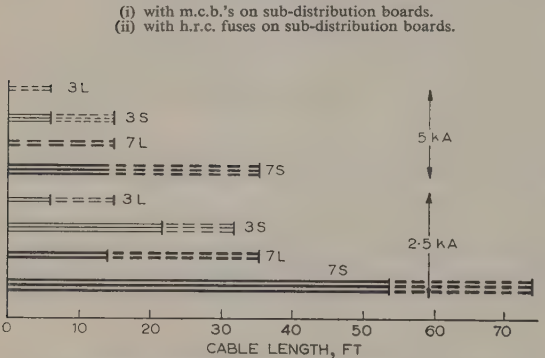


Fig. 2.—Minimum cable length for limiting faults to 5 kA and 2 kA (see Table 3 in Reference 1).

Cable sizes: 3 = 3/·036 in; 7 = 7/·036 in.
Fault types: L = Loop; phase-neutral fault.
S = Severe; 3-phase fault.

(Earth faults depend on earth-return resistance, are always milder than S-type, and probably similar to L-type faults.)

——— Typical installation with 0·03-ohm mains.
– – – Installations with 0·005-ohm mains (severe conditions).

(2.3) M.C.B. Behind H.R.C. Fuse

One of the main conclusions of this analysis is that in many cases it would be functional and economical to use m.c.b.'s on the sub-distribution boards, where their breaking capacity is adequate and where convenient reinstatement of the relatively frequent load-end faults is particularly desirable, and h.r.c. fuses on the main distribution board, where their high breaking capacity is fully exploited and where they compete with relatively expensive oil or air circuit-breakers. This layout leads to interesting problems of discrimination and back-up protection, which can become embarrassing if the choice of equipment is casual or if the radical difference between the functioning of fuses and circuit-breakers is not understood.

(3) DISCRIMINATION AND BACK-UP PROTECTION

(3.1) Requirements

While discrimination over the widest possible current range is desirable with equipment of equal breaking capacity, e.g. h.r.c. fuses on all distribution boards,[4] the requirements of mixed installation are quite different.

If the fault current is within the breaking capacity of th m.c.b.'s on the sub-distribution boards, discrimination is desire and the m.c.b. should clear the fault within the pre-arcing tim of the fuse on the main distribution board. However, if th fault is too severe for the m.c.b., considerations of safety overric the desideratum of discrimination and the fuse in front must a as back-up protection and blow before the circuit-breaker damaged. These two requirements imply that the current/tim characteristics of fuse and circuit-breaker should cross. Fo tunately this is always the case, because of the different breakin mechanisms of the two devices. A circuit-breaker normall interrupts current flow near a current zero, i.e. at the end of half-cycle. For a heavy short-circuit one half-cycle, i.e. 0·01 se can be taken as a nominal breaking time (see Fig. 3), althoug the actual time of current flow depends on the instant of fau initiation, which introduces a random effect. An h.r.c. fus on the other hand, is able to cut off a large current in a fractio of a half-cycle.[4] This means that, in the event of a very heav fault, discrimination is impossible but backing-up occurs.

(3.2) Discrimination Criterion

The question arises: at what fault current does the transitio from discrimination to back-up protection take place? Th author has previously proposed as a working hypothesis[1,2] tha the discrimination limit is determined by the 0·01 sec fusin current. Fig. 3 shows the current/time characteristics of a

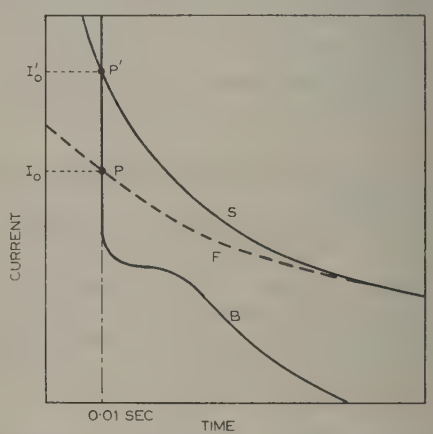


Fig. 3.—Current/time characteristics of slow fuse, S, fast fuse, F and m.c.b., B.

If the fuse is on the main distribution board and the m.c.b. on the sub-distributi board, P (resp. P′) is the take-over point, whereby discrimination occurs for fa currents below I_0 (resp. I'_0) and back-up protection takes place for faults above (resp. I'_0).

m.c.b., a fast fuse and a slow fuse. The take-over point (respectively P′) on the intersection of fuse characteristic and th vertical line for $t = 0{\cdot}01$ sec indicates the discrimination limi one should expect discrimination if the fault current is below I while for larger currents the fuse acts as back-up protection t the circuit-breaker. Take-over is in practice almost independe of the m.c.b.'s normal current rating, but determined by th choice of the fuse.

(4) EXPERIMENTS

(4.1) Scope and Method of Tests

In order to verify the criterion postulated above, discrimina tion tests were carried out on two series of h.r.c. fuses of differe

make in the range between 30 and 160 amp (for previous tests see Reference 1). According to the manufacturers, one of these fuse types should be fast and the other should be relatively slow, although both are faster than semi-enclosed fuses (see Fig. 7). Two makes of m.c.b.'s were used, each of 5, 15 and 30 amp current rating, their rated breaking capacity being 2·5 kA. Currents were measured by means of a cathode-ray oscillograph and camera, which thus gave a record of both magnitude and waveshape. (For method of experiments see Section 10.2.) All oscillograms (except initial tests made at low currents for calibration purposes) consisted of one or two loops only. They can be classified into six groups (Fig. 4) and are discussed in Section 10.2.

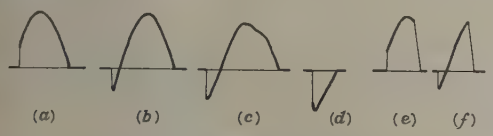


Fig. 4.—Oscillogram types.

(*a*) Full loop (containing the current peak).
(*b*) Full loop, with minor loop.
(*c*) Stunted loop, with minor loop.
(*d*) Tail-end.
(*e*) Cut-off (single loop, post-peak).
(*f*) Cut-off, with minor loop.

(4.2) Results

The discrimination results are presented in Figs. 5 and 6. R.M.S. short-circuit currents, i.e. 0·71 of the observed peak, are plotted against fuse rating; discrimination is indicated by a cross, non-discrimination by a circle. In spite of occasional gaps

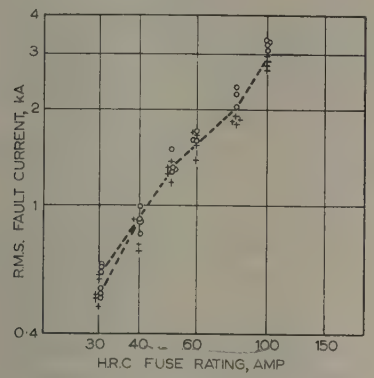


Fig. 5.—Test results on slow fuses.

Tests where the fuse blew are marked by circles; tests where discrimination was obtained by crosses.
R.M.S. fault currents are defined as 0·71 peak values, scaled from oscillograms. Only some of the highest discrimination values and the lowest non-discrimination values are shown; numerous crosses below and several circles higher up were omitted; 136 tests were made on 35 fuses.

or random overlap this limit of the discrimination region can be easily ascertained and is shown by a line. The two lines are redrawn in Fig. 7, which shows for further comparison the 0·01 sec fusing currents of the two fuse types tested as taken from makers' publications, and a previously established line for semi-enclosed fuses.

(4.3) Discrimination Between Fuses and Circuit-Breakers

The tests confirmed that discrimination depends on the fault current: it was always obtained at values considerably below the limit lines in Figs. 5 and 6, and never at currents far above it. It was also confirmed that the discrimination limit depends

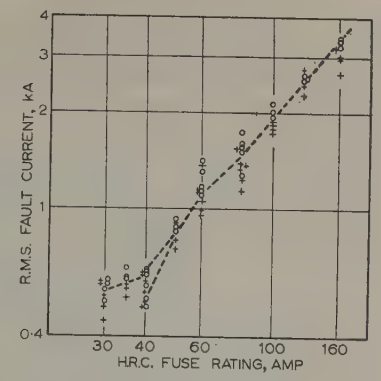


Fig. 6.—Test results on fast fuses.

Symbols and comments as for Fig. 5. 170 tests were made on 59 fuses.

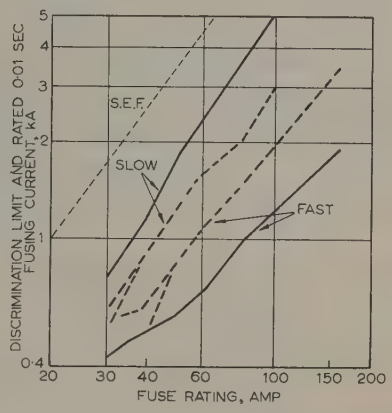


Fig. 7.—Discrimination limit and 0·01 sec fusing current.

– – – – Limits, redrawn from Figs. 5 and 6.
——— 0·01 sec fusing currents, taken from h.r.c. fuse characteristics published by makers.
..... Semi-enclosed fuses, redrawn from Fig. 6 of Reference 1.

on fuse rating, but not significantly on the m.c.b. rating [except that with 30 amp fuses, 5 amp circuit-breakers discriminated better than 30 amp ones in the reduced-voltage tests with traces of type C (Fig. 4) which is not important in practice]. In many practical cases, of course, the relatively high coil resistance of m.c.b.'s of small normal current rating can contribute noticeably to the circuit impedance, and hence can promote discrimination indirectly by limiting the fault current.

The measured discrimination limits did not agree very closely with published 0·01 sec values, being above the published line for the faster fuses and below for the slower ones (Fig. 7). The opposite direction of the discrepancies speaks against a modification of the 0·01 sec rule. These discrepancies cannot easily be explained without a knowledge of the test methods by which different manufacturers obtain their fuse characteristics, and might well be due to different test power factors or different interpretations of statistical results.

Although the difference between the two fuse types tested is less than expected, it is consistent and large enough to be significant: the discrimination limit of the slow fuses lies about 50% above that of the fast fuses (see Fig. 7). It appears that published characteristics can give a useful lead to the choice of fuses, but do not permit accurate discrimination design.

(4.4) Observations on M.C.B.'s

An analysis of minor loops in the test oscillograms suggests that the m.c.b.'s tested usually clear before the next current zero if a fault occurs at least 60° (i.e. 0·003 sec) before this current zero, although occasionally minor loops in the order of 0·004 sec were observed.

For the higher current tests a 100 amp air circuit-breaker of 10 kA rated breaking capacity was connected in series with the test m.c.b. and fuse. One can conclude from an observation of the flash emanating from the smaller test circuit-breaker that the larger one is quite able to provide back-up protection by sharing the breaking duty, although at the expense of discrimination. According to the published characteristics one could not expect the 100 amp circuit-breaker to discriminate against a smaller one for short-circuit currents above about 1 kA.

Two of the 30 amp circuit-breakers finally failed to operate, owing to burnt contacts, after having cleared several short-circuits above, and numerous short-circuits just within, their rated breaking capacity.

(5) STATISTICAL THEORY OF DISCRIMINATION

The random effect mentioned in Sections 3 and 4 will now be scrutinized, and an attempt will be made to explain the statistical nature of discrimination tests.

Curve H in Fig. 8 shows the heat produced in a conductor by

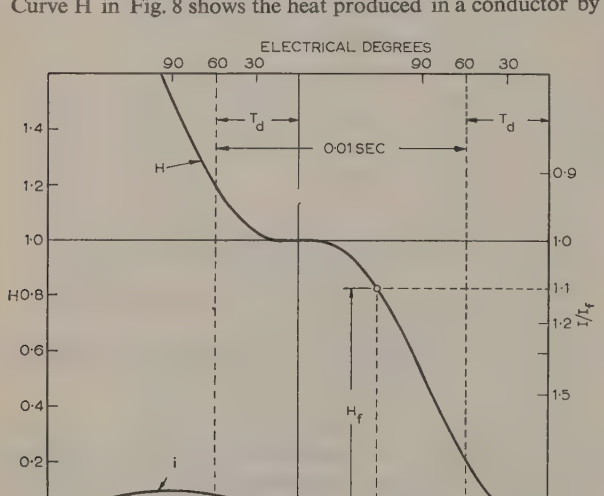


Fig. 8.—Calculated curve for heat developed between time t_0 and current zero at time $2\pi/\omega$.

Differences of H-curve show heat between any two instants.
Phase is indicated by i-curve. Inverse-square-root-scale on right correlates H_f, the 0·01 sec melting heat of fuse, with H-curve. Electrical degree scale on top refers to circuit-breaker-delay, T_d.

a sinusoidal current suddenly starting at time t_0 and ending when current is zero at time $2\pi/\omega$. (See Section 10.3 for proofs.) Let us assume that a particular fuse will require heat H_f to blow, and that an m.c.b. in series with that fuse has a delay time T_d, which means that all faults initiated between t_1 and t_2 will be cleared at $2\pi/\omega$. Since there is an equal chance of the fault occurring at any moment in the cycle, the ratio $\omega(t_2 - t')/\pi$ is the discrimination probability for this particular circuit-breaker, fuse and current (see Section 10.3).

Fig. 9(*a*) shows the discrimination probability as a function of current, expressed in terms of the 0·01 sec fusing current, I_f,

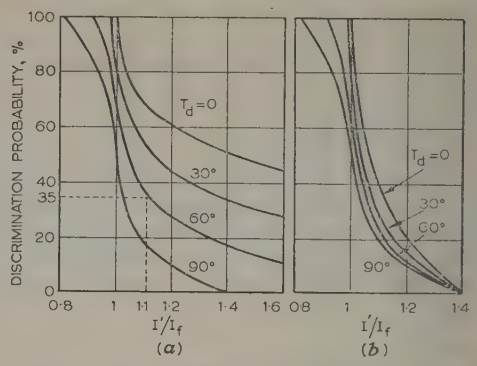


Fig. 9.—Discrimination probability as function of current with circuit-breaker delay as parameter.

Curves were obtained from Fig. 8. Abscissa in Fig. 9(*a*): Prospective fault current in terms of I_f.
Abscissa in Fig. 9(*b*): Observed fault current in terms of I_f.

of a certain fuse, with different circuit-breaker delay times a parameter. As an example of how to read the diagram, a short circuit current of 1·1 I_f has a 35% probability of discriminatio if the delay is 60°. Fig. 9(*a*) refers to prospective currents; i tail-end clearances are excluded, Fig. 9(*b*) is obtained, whic shows the discrimination probability for currents as observe on test oscillograms (see Section 10.3).

Fig. 9(*a*) allows the conclusion that the 0·01 sec fusing current criterion is quite practicable, since the drop in the regio of unity is fairly sharp; it is somewhat conservative for moderat circuit-breaker delay, since discrimination can occasionally occu at very much higher currents, although it is only fully certai at a slightly lower current. Fig. 9(*b*) explains why the result of discrimination tests are bound to be 'statistical', even whe based on current peaks. The curves are so steep, however that a discrimination limit can be ascertained without a extravagant expenditure of time and fuses.

The minor-loop analysis in Section 4.4 indicates that the 60 curve is applicable to our tests.

(6) PRACTICAL CONCLUSIONS

(6.1) Choice of Fuses for Use in Front of M.C.B.'s

The use of h.r.c. fuses on a main distribution boar with m.c.b.'s on sub-distribution boards is functional, an should allow satisfactory operation if a suitable take-over poin between discrimination and back-up protection is chosen (se Fig. 3). The choice depends on the rated breaking capacity o the m.c.b. The author would favour a take-over range betwee 40 and 100%; this means that the fuse must take over belo the breaking-capacity current of the m.c.b., but should not d so below say 40% of this value. The top limit is a matter o safety; the lower limit prevents unnecessary loss of discrimina tion and is fortunately not critical.

If this tolerance range is applied to the fuses tested, Fig. shows that fast fuses are undesirable in circuits below 80 amp since they lose discrimination already in the region of 1 kA In circuits, carrying more than 80 amp, slow fuses would no provide back-up protection for a 2·5 kA circuit-breaker. 5 kA circuit-breakers would combine well with slow fuses for circuit carrying between 60 and 150 amp, or with fast fuses for circuit carrying 80–200 amp.

(6.2) Suggestion of New Fuse Series

It appears that neither series of h.r.c. fuses tested is ideal fo co-ordination with m.c.b.'s. This could not be expected, sinc

existing fuse series have been devised for optimizing installations in which fuses are used exclusively.[4]

What then would be the ideal characteristics of a fuse series for a main distribution board, if m.c.b.'s are to be used on sub-distribution boards?

Two requirements should be met, namely

(*a*) Correct overload protection of sub-main cables between main- and sub-distribution boards; this determines the fuse rating.
(*b*) A suitable take-over point from m.c.b.'s on the sub-distribution boards, determined by the breaking capacity of the m.c.b.'s.

Fig. 10 shows two sets of possible characteristics, for m.c.b.'s with breaking capacities of 2·5 and 5 kA respectively, with take-

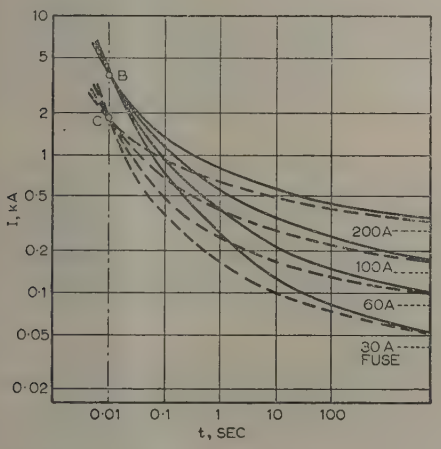


Fig. 10.—Ideal characteristics for main distribution board fuse series for backing up m.c.b.'s on sub-distribution boards.

——— Circuit-breakers of 5 kA breaking capacity.
– – – – Circuit-breakers of 2·5 kA capacity.
I = Prospective current, kA.
t = Pre-arcing time, sec.

over points placed at 75% to allow safely for tolerance in characteristics. These curves present an ideal which would allow quite a wide tolerance in execution. Whether fuse manufacturers can provide a fuse series anywhere near the ideal is a technological and economic problem beyond the scope of the paper. The main requirement seems to be a slowing down of the smaller fuses. It might be pointed out that a sacrifice in breaking capacity thereby entailed would be quite acceptable in most practical applications, since the distribution-board fuse has rarely to cope with currents much above 10 kA and can in turn be backed up by mains protection of higher rating if necessary.

(6.3) Back-Up by Air Circuit-Breakers

Since relatively low-priced air circuit-breakers have become available in the 50–100 amp range and with breaking capacities in the region of 10–20 kA, the exclusive use of circuit-breakers has been found economical. Back-up protection and discrimination are still a problem, however. Fast air circuit-breakers can lighten the breaking duties of the m.c.b. behind considerably for heavy faults above a certain current, at the expense of discrimination of course. It was not the task of the tests performed to establish this limit carefully. Published characteristics lead to the conclusion that discrimination is lost at current levels comparable with h.r.c. fuses. Although the inconvenience thus caused is mitigated by the easy reinstatement of circuit-breakers, manufacturers might consider the design of a series with higher take-over independent of normal current rating, on principles similar to those suggested in Section 6.2.

(6.4) Oil Circuit-Breakers

Most oil circuit-breakers are too slow to provide effective back-up protection to an m.c.b.[1] This means the circuit-breakers or fuses on the sub-distribution boards must have a breaking capacity to clear faults at their own terminals. There is no discrimination problem with oil circuit-breakers on the main distribution board, but they come into play very rarely, and their high capital cost seems to bear small dividends.

(6.5) Mixed Distribution Boards

A number of mixed layouts have been suggested before.[1] To summarize briefly:

(*a*) Mixed sub-distribution boards: m.c.b.'s for long load circuits and h.r.c. fuses for short runs (see Figs. 1 and 2).
(*b*) Mixed main distribution boards: a mixture of slow h.r.c. fuses for circuits up to 80 amp and fast h.r.c. fuses for heavier circuits would give a design based on fuses at present available, which is comparable with, though not quite as good as, the fuse series proposed in Section 6.2.
(*c*) Double protection by m.c.b.'s and back-up h.r.c. fuses, both on the main distribution board (with only m.c.b.'s on the sub-distribution boards): this promises excellent performance at a slightly higher price.

(6.6) Protection Policies

It is a matter of policy whether protection devices on sub-distribution boards are chosen to cope with faults at their own terminals, e.g. h.r.c. fuses [see Fig. 1(ii)], or whether reliance is placed on back-up protection on the main distribution board for this infrequent task, clearing only load-end faults at the sub-distribution board, e.g. an m.c.b. backed-up by a h.r.c. fuse [see Fig. 1(i)].

With correct take-over design the latter solution implies only a loss of discrimination, and not safety. The decision rests between the convenience of the m.c.b. but the loss of larger sections of the installation in cases of heavy faults, and better discrimination, but the nuisance and cost of relatively frequent fuse replacements for load-end faults.

(7) SUMMING-UP

Miniature circuit-breakers are convenient on sub-distribution boards and capable of clearing most load-end faults; busbar faults usually exceed their breaking capacity in large installations, and h.r.c. fuses or circuit-breakers of relatively high breaking capacity are then required on main distribution boards.

The use of m.c.b.'s on sub-distribution boards with h.r.c. fuses on the main distribution board is logical and economical, but requires careful discrimination design, so that the fuse will cope with heavy short-circuits while the m.c.b. will clear overloads or moderate faults and thereby achieve discrimination in most cases.

The 0·01 sec fusing current can be considered as the discrimination limit, although a random variation due to the phase instant of fault initiation should often result in fortuitous discrimination at higher currents.

Of the h.r.c. fuses tested, the fast type discriminates poorly for current ratings below 80 amp, while the slow type when rated above 80 amp will not provide adequate back-up protection for 2·5 kA circuit-breakers. Characteristics of a series of h.r.c. fuses ideal for m.c.b. back-up protection are suggested in Fig. 10.

Rapid-acting air circuit-breakers can provide back-up protection for m.c.b.'s, but their characteristics must be carefully chosen lest discrimination be lost with relatively mild faults.

Mixed distribution-board equipment may sometimes be warranted. Double main-distribution-board protection by fuses and circuit-breakers together offers excellent design possibilities.

The author is convinced that both h.r.c. fuses and m.c.b.'s have their place in low-voltage installations. The paper is intended to aid in the choice and thus contribute to rational design.

(8) ACKNOWLEDGMENTS

The recent tests at the University of Cape Town were carried out by R. D. Lockyear, some jointly with the author, others under his supervision. The efforts and co-operation of Mr. Lockyear are gratefully acknowledged.

Thanks for samples and trade information are due to the following companies and their Cape Town representatives: Hubert Davies and Co. for Murray and Heinemann, Ltd., the English Electric Co., Ltd., F. W. J. Electrical Industries Ltd., and Reyrolle and Co., Ltd.

(9) REFERENCES

(1) EINHORN, H. D.: 'The Protection of Low-Voltage Installations by means of Miniature Circuit-Breakers and H.R.C. Fuses', *Transactions of the South African Institute of Electrical Engineers*, 1957, **48**, p. 3.

(2) EINHORN, H. D.: Discussion on 'Design, Performance and Application of Miniature Circuit-Breakers', *Proceedings I.E.E.*, 1955, **102**, p. 373.

(3) MIDDLECOTE, A. A.: Discussion to Reference (1), p. 22.

(4) JACKS, E.: 'Discrimination between H.R.C. Fuses', *Proceedings I.E.E.*, Paper No. 2805U, January, 1959 (**106** A, p. 299).

(10) APPENDICES

(10.1) Fast and Slow H.R.C. Fuses

'Fast' and 'slow' h.r.c. fuses are meant to be relatively fast or slow in the 0·01 sec region. It is quite possible that a slow fuse may have a smaller fusing factor than a fast one, which would imply crossing of the two fuse characteristics in Fig. 3. A quantitative definition of 'fast' and 'slow', if wanted, can be obtained from the relation:

$$I_{0\cdot01} = kI_r^{4/3}$$

One could, for example, agree to call fuses 'slow' if $k > 7$, 'fast' if $k < 4$, and 'medium' if $5 < k < 6$. With this definition, the two fuse types tested (see Figs. 5–7) would form border-line cases. For them, as probably for most conventional fuse series, k is substantially constant (between 4 and 4·5 for the faster, between 6 and 7 for the slower, types).

Table 1, based on the above definitions, might aid in installation design.

Table 1

RECOMMENDATION FOR CHOICE OF FUSE SPEED

Fuse rating, amp	30	40	50	60	80	100	125	150	200
For 2·5 kA m.c.b.	vs	s	s, m	s, m, f	s, m, f	m, f	f	vf	vf
For 5 kA m.c.b. ..	vs	vs	s	s	s, m	s, m, f	s, m, f	m, f	f

s = Slow.
m = Medium.
f = Fast.
vs and vf = Well within the slow or fast region.

(10.2) Experimental Method

A short-circuit was applied to the test h.r.c. fuse, the test m.c.b. and a 4-terminal resistor of 0·01 ohm (resp. 0·05 ohm for smaller currents), all connected in series (Fig. 11). The voltage across the 0·01-ohm resistor was applied to a cathode-ray oscillograph, for recording the current magnitude and waveshape photographically.

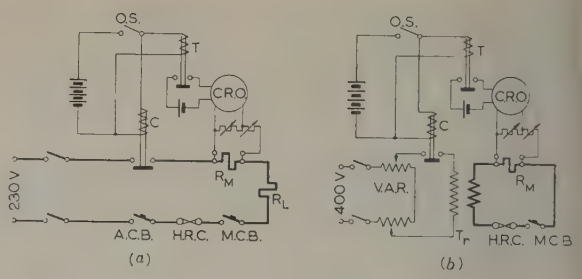


Fig. 11.—Test circuits.

(*a*) For currents exceeding 0·8 kA, directly from substation busbars.
(*b*) For currents less than 1·2 kA, via 230/18-volt transformer.
OS = Operating switch.
T = Time-sweep relay.
C = Current-closing relay.
R_L = Current-limiting resistor.
R_M = Four-terminal measuring resistor.

Tests up to about 1·2 kA were performed in the laboratory over a 230/18-volt 2 kVA transformer; the current was controlled by two variable transformers in series from 380 volts [Fig. 11(*b*)]. Tests between 800 amp and 3·5 kA were performed in the 600 kVA substation, directly from a spare oil circuit-breaker on the 230-volt main distribution busbars; the current was adjusted by limiting resistors of copper wire [R_L in Fig. 11(*a*)].

The oscillograph time sweep through relay T, and the closing of the short-circuit current through the suitably delayed relay C were both initiated by the hand-operated switch O.S.

The current scale was adjusted by a decade box potentiometer and precalibrated by means of ammeter and current transformer at lower currents. Extrapolation of current calibration to very much higher values was unavoidable, but the accuracy thus achieved was adequate for short-circuit tests where random effects enter and high accuracy is not needed.

The currents recorded in Figs. 5 and 6 were based on the peaks observed. Fig. 4 shows that, in cases (*a*) and (*b*), but not in cases (*c*) and (*d*), they correspond to the short-circuit currents as determined by the network impedances which, in fuse-engineering parlance, are called 'prospective' short-circuit currents [although the term 'retrospective' would seem to suit the tail-end case (*d*) better]. In cases (*a*) and (*e*) the fault occurs at an instant of rising current, while in the other cases it occurs when the current falls, often resulting in a 'minor loop'. The occurrence of (*b*), (*c*) or (*d*) depends on fault instance, on circuit-breaker speed and on current; in the infrequent case (*c*) the arc resistance curtails the current peak. In case (*d*) discrimination was usually achieved. Cases (*e*) and (*f*) are typical for fuse operation. Cut-off after the peak, as shown in (*e*) was usually observed in our tests; cut-off before the peak, as shown in (*f*) and in fuse publications,[4] is typical for currents considerably above the discrimination limit.

(10.3) Statistical Analysis

In Fig. 8, if a sinusoidal fault current, $i = I\sqrt{2}\sin\omega t$, starts at fault instant t_0 and is terminated by a circuit-breaker at a subsequent current zero at time $2\pi/\omega$, the heat developed in a fuse element of resistance R is

$$H = R\int_{t_0}^{2\pi/\omega} i^2 dt = 2RI^2\int \sin^2 \omega t dt$$

$$= RI^2[1/f - t_0 + (1/2\omega)\sin 2\omega t_0]$$

The heat, H_f, required to blow a particular fuse is a constant

n case of a heavy fault, since the time is too short for any appreciable loss of heat.

Discrimination will occur if $H < H_f$, and the fuse will blow f $H > H_f$. The H-function shown in Fig. 8 is therefore closely elated to the discrimination limit.

The probability of discrimination can now be reasoned out as follows:

There is an equal chance for the fault incidence (t_0) at any instant in the cycle. Let us normalize the H-scale by taking the heat from a particular fault current in half a cycle as unity. If then $H = H_f$ for fault occurring at time t', discrimination occurs when $t_0 > t'$, but the fuse will blow when $t_0 < t'$ (see Fig. 8).

If the circuit-breaker delay time is T_d (which can be expressed in seconds or in electrical degrees) and $t_1 = \pi/\omega - T_d$, $t_2 = 2\pi/\omega - T_d$, the discrimination region is between t' and t_2, and the non-discrimination region between t_1 and t'. The discrimination probability is then

$$\frac{(t_2 - t')}{(t_2 - t_1)} = \frac{\omega(t_2 - t')}{\pi}$$

and can be obtained graphically from Fig. 8 with adequate accuracy. Fig. 9(*a*) shows the variation of discrimination probability with current. The abscissa is obtained by defining $I/I_f = \sqrt{(H/H_f)}$, where I is the fault current and I_f the 0·01 sec fusing current (see right-hand scale in Fig. 8). The current I in Fig. 9(*a*) is a prospective fault current, and corresponds to the actual observed current in cases (*a*) and (*b*) of Fig. 4, but not in case (*d*). The discrimination probability of actually observed currents [ignoring the 'stunted' current case (*c*)] is given by

$$\frac{1{\cdot}5\pi/\omega - t'}{0{\cdot}5\pi/\omega + T_d}$$

since tail-end short-circuits [case (*d*)] are counted in a lower current category.

One can also deduce from the H-curve in Fig. 8 how much the instant of fault incidence would affect the current/time characteristics of fuses. It can be seen that this variation, although zero for the 0·01 sec point itself, can be quite large in its vicinity and that different interpretation of statistical results can at least partly explain the discrepancies in Fig. 7 between the test results and curves culled from published fuse characteristics.

It should be noted that this statistical analysis, including Figs. 8 and 9, is based on the assumption of non-inductive circuits, which means that the short-circuit current can jump at instant t_0 to a value i which is entirely limited by resistance. This assumption applies with good approximation to our tests and in practice to faults where the current is limited mainly by cables.

621.376.223: 621.374.32: 681.142

The Institution of Electrical Engine
Monograph No. 346
Oct. 19

THE USE OF SILICON DIODES IN D.C. MODULATORS AND THEIR APPLICATION TO DRIFT CORRECTORS FOR COMPUTING AMPLIFIERS

By T. GLUCHAROFF, M.E., and C. P. GILBERT, M.Sc., Associate Member.

(*The paper was first received* 17*th October*, 1958, *and in revised form* 15*th June*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in October*, 1959.)

SUMMARY

The basic computing element of an electronic analogue computer is the high-gain d.c. amplifier, but in order to overcome its inherent drift some form of auxiliary drift-correcting amplifier is frequently used.

The paper describes a silicon diode modulator for use in such drift correctors, and shows that the zero stability which can be achieved is comparable with that given by the conventional relay modulator. It is also shown that the use of the diode circuit results in a number of improvements in the overall drift-corrector performance, such as a higher frequency response and practically unlimited life; the design of such a drift corrector for use with a given d.c. amplifier is described in detail and the resulting performance is assessed.

LIST OF SYMBOLS

a = Resistance ratio of the modulator balancing arm.
b = Diode branch resistance ratio during conduction.
c = Diode branch resistance ratio during non-conduction.
C_p = Parallel combination of the capacitance across the diodes.
V = Amplitude of the switching voltage.
i = Diode current.
i_g = Grid current.
I_L = Diode leakage current.
I_s = Diode saturation current.
$K_1G_1(j\omega)$ = Frequency-response function of the d.c. amplifier.
$K_2G_2(j\omega)$ = Frequency-response function of the drift corrector.
$KG(j\omega)$ = Frequency-response function of the drift-corrected amplifier.
k = A constant.
p = Laplacian operator.
R_0 = Amplifier or drift corrector input impedance when purely resistive.
R_p = Parallel combination of the diode back-resistances.
R_T = Parallel combination of R_0 and R_p.
T = Absolute temperature.
v = Circuit input voltage, or a voltage injected from a supply source.
v' = Amplifier input voltage.
v'' = Amplifier output voltage.
v_i = Modulator input voltage.
v_f = Drift voltage.
v_o = Modulator output voltage.
v_{oc} = Modulator output voltage due to unbalance in the conducting direction.
v_{on} = Modulator output voltage due to unbalance in the non-conducting direction.
v_t = Modulator output voltage due to unbalance in the non-conducting direction in the presence of capacitance.
Z_f = Feedback element of the computing component this may be purely resistive (R_f) or pure capacitive (C_f).
Z_i = Parallel combination of all the input elements.
Z_0 = Amplifier or drift corrector input impedance.
ω_c = Carrier angular frequency.
ω_s = Signal angular frequency.

(1) INTRODUCTION

(1.1) General

The usual method of drift correction used in electron analogue computers has been described by Goldberg.[1] modulator converts the slow variations of the input signal to a alternating voltage; the latter is in the form of a modulate carrier which, after passing through an a.c. amplifier, is demodu lated to obtain an amplified version of the original signal. Suc amplifiers are often referred to as chopper amplifiers, the zer stability obtainable being many times better than that of th best d.c. amplifier.

Drift-correcting amplifiers reduce the drift in a computin amplifier using the circuit shown in Fig. 1(*a*) either with or with

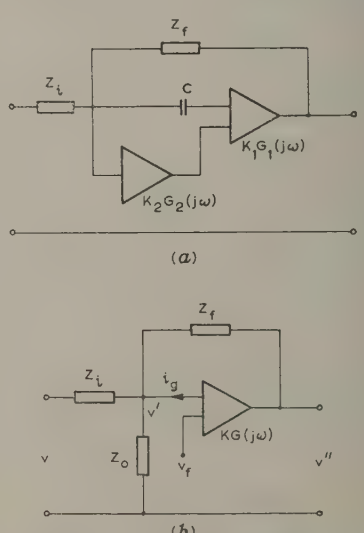


Fig. 1.—Computing amplifier.

(*a*) Basic method of connecting a drift-correcting amplifier: C eliminates the effe of grid current arising in the d.c. amplifier.
(*b*) Drift-corrected computing amplifier complete with its computing component

out an input capacitor.[2] If K_1 and K_2 are the low-frequency gai of the d.c. amplifier and the drift corrector respectively, th low-frequency gain of the combined amplifier becomes K $(1 + K_2)K_1$; the high-frequency response is the same as that the d.c. amplifier alone. Also, the reduction of drift of the d.

Correspondence on Monographs is invited for consideration with a view to publication.
Mr. Glucharoff is Lecturer, and Mr. Gilbert is Senior Lecturer, in Electrical Engineering, University of New South Wales.

mplifier is proportional to $(1 + K_2)$, and if K_2 is sufficiently irge the drift of the drift corrector alone will be of importance. n general, drift correctors reduce the input resistance of the ombined amplifier and increase the noise level considerably. n addition, grid-current effects are not reduced by the drift orrector unless an input capacitor is used as shown in Fig. 1(*a*). n the following, a short description of the various amplifier mperfections is given, and the main effects of the drift corrector re briefly examined.

(1.2) Drift Corrector Requirements

Consider a drift-corrected computing amplifier as shown in Fig. 1(*b*), where Z_i and Z_f are the computing components, Z_0 is he input impedance, K is the low-frequency gain of the combined amplifier, i_g is the grid current and v_f is the drift voltage eferred to the input. Then

$$\frac{v - v'}{Z_i} = \frac{v' - v''}{Z_f} + \frac{v'}{Z_0} \pm i_g$$

nd since $$v' \pm v_f = -v''/K,$$

$$\frac{Z_f}{Z_i} = -v'' - \frac{v''}{K}\left(1 + \frac{Z_f}{Z_i} + \frac{Z_f}{Z_0}\right) \pm v_f\left(1 + \frac{Z_f}{Z_i} + \frac{Z_f}{Z_0}\right) \pm i_g Z_f \quad . \quad . \quad . \quad . \quad (1)$$

The second, third and fourth terms on the right are the errors due to finite gain, drift and grid current respectively.

The error due to finite gain can be made as small as desired by making K sufficiently large, but it should be noted that K is the low-frequency gain only, and at higher frequencies the gain will be less, causing larger errors. To keep these errors small, a good high-frequency response is required.

The error due to drift is $v_f(1 + Z_f/Z_i + Z_f/Z_0)$, and is independent of K. In an integrator with $Z_i = R_i = 1$ megohm, $Z_0 = R_0 = 1$ megohm and Z_f capacitive with $C_f = 1\,\mu$F, the error due to drift is

$$v_f + \frac{R_i + R_0}{R_i R_0}\,\frac{1}{C_f}\int v_f\,dt$$

If $v_f = 10$ mV, this output term will change at the rate of 20 mV/s and if a maximum output of 100 volts and an accuracy of 0·1% is required, the integrating time must be limited to only 5 sec. In practice, errors of this type are usually far larger than any others, and this is where the main improvement in accuracy can be obtained by the use of drift correctors.

The output due to grid current is $i_g Z_f$ and in an integrator the resulting error becomes $i_g dt/C_f$. For $i_g = 10^{-9}$ amp and $C_f = 1\,\mu$F, a rate of change of output of 1 mV/s is present and this may itself limit the integration time for a given accuracy. Clearly, in order to obtain the full benefit of reduced drift, the grid current must be kept low. Electrometer valves are normally employed, giving grid currents of less than 10^{-10} amp, but a more effective method of reducing grid current effects is to use a capacitive input to the d.c. amplifier, Fig. 1(*a*).

The amplifier input impedance, Z_0, appears in the finite gain and drift error terms but always in parallel with the input impedance, Z_i. The latter seldom exceeds 1 megohm, and so Z_0 should be at least as large as this to ensure that its contribution to these error terms is negligible.

The noise voltage is increased by the use of drift correctors and may be as large as 5 mV referred to the input. Normally it has little effect upon adders or integrators, the latter tending to cause a reduction in noise level. In differentiators, however, the reverse is true, and a very large noise output may result. For this reason it is desirable to keep the combined amplifier noise level to a minimum.

The above discussion indicates that the desirable features of a drift corrector are low drift, high input impedance, low noise and good high-frequency response.

The advantages of a good high-frequency response are:

(*a*) The drift-corrector gain increases, not only the accuracy of integration, but also that of most other computing operations.

(*b*) The size of the series capacitor used to remove grid current effects [Fig. 1(*a*)] can be reduced to reasonable proportions; the amplifier input time-constant can now be kept small, thus removing the main objection to this type of circuit.

(*c*) In general, a higher carrier frequency will permit considerable reduction in the physical size of some of the components used.

The best known type of drift corrector uses a mechanical vibrator operating at some frequency below 400 c/s. Zero stability of 0·1 mV or better can be obtained during its life, which, however, may be limited.

In the paper, a silicon-diode modulator is described which is suitable for use in drift-correcting amplifiers for computing applications. Its performance is comparable with the mechanical modulator in most respects, but it has the advantage of higher switching frequency, better frequency response and practically unlimited life. This unit has been designed to work with the d.c. amplifier used in Utac [University of New South Wales (formerly University of Technology) Analogue Computer], but is suitable for much wider application with only slight modifications.

(2) THE DIODE MODULATOR

If two diodes are connected as shown in Fig. 2(*a*) and an alternating voltage is applied, the diodes will alternately conduct

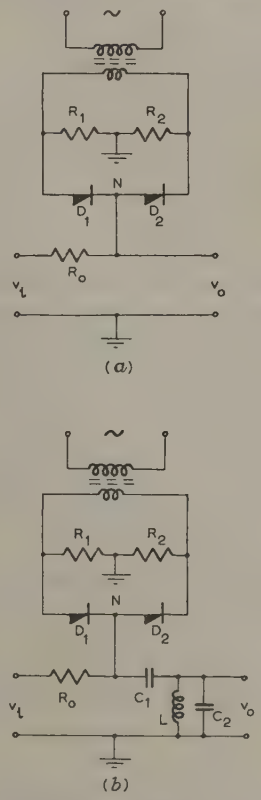


Fig. 2.—Modulator.

(*a*) Basic circuit.
(*b*) Modulator working into a tuned circuit.

or block. The switching voltage is balanced to earth by two low resistances, R_1 and R_2, so that, if the characteristics of the diodes are identical over the whole operating region, their junction point, N, will remain at zero potential at every instant. If a direct voltage, v_i, is applied to the junction point via a resistor R_0, then, provided that R_0 is much larger than the forward resistance and much smaller than the backward resistance of the diodes, the voltage v_o at the junction point will be a square wave alternating between zero and v_i. Thus, a direct voltage is converted into a pulsating voltage which can be amplified by a conventional a.c. amplifier.

For an input voltage $v_i(t)$ the output voltage, $v_o(t)$, is given by

$$v_o(t) = v_i(t)\left[\frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin(2n-1)\omega_c t}{2n-1}\right]$$

where ω_c is the angular frequency of the switching voltage and the second term in the brackets represents the square wave. With a sinusoidal input, $v_i = A\sin\omega_s t$, for which ω_s is much smaller than ω_c, the first term in the brackets can be neglected since it is filtered out in the a.c. amplifier. Then by considering the fundamental term only, the output becomes

$$v_o(t) = \frac{2}{\pi}A\sin\omega_s t\sin\omega_c t$$

which is the equation of a suppressed-carrier amplitude-modulated wave; this is fed into the a.c. amplifier.

However, if stray capacitance and capacitance required for modulator balancing are considered, the total input capacitance of the a.c. amplifier may become as high as 50 pF, and this in series with R_0 will give severe attenuation and phase shift at the switching frequency of interest.

To eliminate this effect, an inductance, L, may be used to tune the circuit, as shown in Fig. 2(*b*), in which C_1 is just large enough to pass the carrier frequency. For satisfactory operation, the impedance of the tuned circuit at resonance should be several times larger than the input resistance R_0. In this case the suppressed-carrier amplitude-modulated wave is applied to a resonant circuit tuned to the carrier frequency. The effect of such a circuit upon the signal frequency[3,4] is the same as that of a single RC network whose break frequency is $1/2\pi R2C$, where C is the total capacitance of the resonant circuit and R is the total equivalent resistance; above this frequency the signal voltage will be attenuated at 6 dB per octave. In the present application, an additional 250 pF is connected at the amplifier input, making the break frequency about 300 c/s. The use of this larger capacitance has the advantage of rendering the circuit less dependent upon inter-electrode and stray capacitances, which are liable to vary and affect the modulator balance.

The tuned circuit also acts as a band-pass filter; it tends to reduce any components of the input signal which are not superimposed on the carrier, higher harmonics, switching transients and power-frequency pick-up.

A switching frequency of 10 kc/s was chosen, since this was already available in the computer Utac. However, it should be pointed out that a higher switching frequency would result in no improvement in this application since, as will be shown later, the gain of the drift corrector must become less than unity at about 200 c/s and this frequency can be easily modulated on a 10 kc/s carrier.

(3) THE SILICON DIODE

In the previous Section it was assumed that identical diodes were used, allowing perfect balance to be obtained at the junction point N. In practice, the characteristics of the diodes will v widely and several adjustments may be required to obtain satisfactory balance.

The d.c. characteristic of a silicon diode for a few tenths o volt in the forward and backward directions is well defined the relation[5,6]

$$i = I_s(\varepsilon^{v/kT} - 1) + I_L\left(\varepsilon^{v/kT} - \frac{v}{kT} - 1\right)^{1/2} \quad . \quad .$$

where I_s is the reverse saturation current and v is the volta across the diode.

A typical semilogarithmic plot in the forward direction shown in Fig. 3. Examination of the diode equation shows th there are three parameters to be considered, I_s, I_L and k.

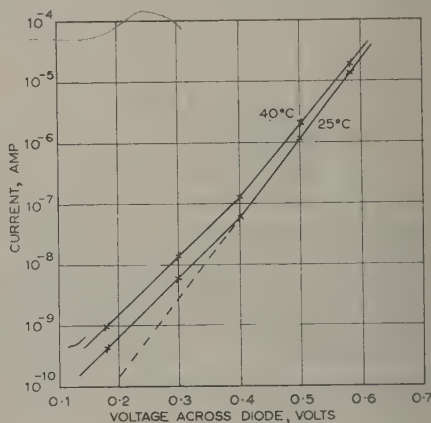


Fig. 3.—Forward characteristic of a silicon diode.

——— Indicating the effect of a temperature change.
– – – Type of characteristic obtained for very low diode leakage currents.

I_s varies considerably from diode to diode but may be a small as 10^{-14} amp. Its effect is to give logarithmic behaviou of the diode up to 0·6 volt in the forward direction; above th voltage large deviations from the logarithmic law can b expected.

I_L is of the order of 10^{-6}–10^{-12} amp and will determine th current in the forward direction up to the point where saturatio and leakage-current effects become comparable. In Fig. this is shown as a change in slope of the logarithmic charac teristic to roughly half of the original value below 0·4 volt (fu line). In the reverse direction, the total current is determined b I_L, the effect of I_s being negligible. Diodes with low leakag current are characterized by excellent logarithmic characteristi in the range +0·1–+0·5 volt (dotted line) and very hig resistance in the non-conducting direction.

The constant k determines the slope of the logarithmic curv and as a rule does not vary widely for a given make of diode.

Eqn. (2) also shows that the diode current is a logarithmi function of temperature. This temperature effect appears t differ only slightly for all the diodes examined, and typical curve are given in Fig. 3 for two temperatures.

The a.c. performance of semiconducting diodes is usuall characterized by their recovery time, rectification efficiency an capacitance, but these properties showed no significant effe upon the performance of the modulator at a switching frequenc of 10 kc/s. The capacitance of a silicon diode may be as lo as 1 pF.

(4) BALANCE IN PRACTICAL MODULATORS

(4.1) Low-Frequency Switching Voltage

If a switching voltage of square waveform and low repetition rate is applied to the modulator circuit shown in Fig. 4, in order to obtain zero potential difference between the diode junction N

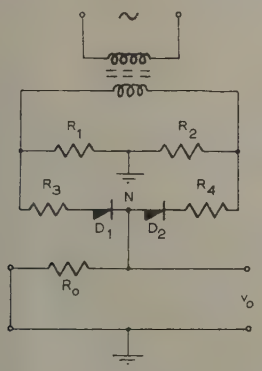


Fig. 4.—Circuit illustrating the sources of modulator drift.

and earth during the conducting period, the required condition is that

$$\frac{R_1}{R_2} = \frac{R_3 + \text{resistance of diode 1}}{R_4 + \text{resistance of diode 2}}, \text{ or } a = b$$

where $a = R_1/R_2$ and $b = (R_3 + \text{resistance of diode 1})/(R_4 + \text{resistance of diode 2})$.

For a sufficiently large switching voltage of amplitude V, the diode resistance can be smaller than R_3 and R_4 and hence b will depend mainly upon these two resistors. If $a \neq b$, an unbalance voltage, v_{oc}, is developed:

$$v_{oc} = V \frac{R_0}{R_0 + \frac{R_1R_2}{R_1+R_2} + \frac{R_3R_4}{R_3+R_4}} \left(\frac{R_2}{R_1+R_2} - \frac{R_4}{R_3+R_4} \right)$$

$$= V\left(\frac{1}{1+a} - \frac{1}{1+b} \right)$$

provided that R_0 is much larger than

$$\frac{R_1R_2}{R_1+R_2} + \frac{R_3R_4}{R_3+R_4}$$

which is normally the case.

During the non-conducting period the condition for zero output at N is that

$$\frac{R_1}{R_2} = \frac{\text{Resistance of diode 1}}{\text{Resistance of diode 2}}, \text{ or } a = c$$

where $c = (\text{Resistance of diode 1})/(\text{Resistance of diode 2})$, since the diode resistances are much larger than R_3 and R_4.

If $a \neq c$, an unbalance voltage, v_{on}, is developed such that

$$v_{on} = V \frac{R_0}{R_0 + R_p} \left(\frac{1}{1+a} - \frac{1}{1+c} \right)$$

where R_p is the parallel combination of the diode resistances in the non-conducting period. The unbalance in this case is largely dependent upon the input resistance R_0.

If the ratio a is made variable, the conditions of interest which may arise are:

(a) $v_{oc} = v_{on}$, i.e. the voltages at N during both conducting and non-conducting periods are the same. For this type of unbalance there is no alternating voltage present at N but only a direct voltage of magnitude $v_{oc} = v_{on}$ which causes a current to flow in R_0 and the signal-voltage source. If, as in computing amplifiers, high impedances are involved, this may be very undesirable.

(b) $v_{oc} = -v_{on}$. In this case a pure alternating voltage appears at N. The main process of balancing the modulator is to reduce this voltage to a minimum; if present, it will give rise to a spurious output and, as will be explained later, may cause drift due to gain changes of the following a.c. amplifier.

(c) $v_{oc} \neq v_{on}$. This condition is a combination of cases (a) and (b) and so both the defects mentioned above will be in evidence to some extent.

The only case in which the potential difference between the diode junction N and earth will remain zero at all times is when $v_{oc} = v_{on} = 0$ and this condition can only be met if $a = b = c$.

Clearly, a control must be introduced to vary either b or c, in addition to that required to vary a, in order to meet this requirement. Complete balance cannot be achieved in practice, and condition (c) normally prevails.

(4.2) High-Frequency Switching Voltage

Consider now a switching voltage of a higher frequency, of either square or sinusoidal waveform. Capacitive and transit effects, i.e. the behaviour of the diodes in the intermediate condition between fully conducting and fully non-conducting, must be considered.

During conduction the condition remains as before ($a = b$).

During the non-conducting period the capacitive reactance may be many times smaller than the resistance of the diodes and so the condition for balance is

$$a = \frac{\text{Impedance of diode 1}}{\text{Impedance of diode 2}}$$

Thus, both the phase and magnitude of the diode branch impedance should be controlled.

To ensure complete balance during the transit period the incremental resistances of the diodes must be equal at every point of their characteristics.[7] From eqn. (2), with I_L neglected, $i = I_s\varepsilon^{v/kT}$ for v slightly positive, and $\frac{di}{dv} = I_s\frac{\varepsilon^{v/kT}}{kT} = \frac{i}{kT}$ or $\frac{dv}{di} = \frac{kT}{i}$, i.e. for every current i and temperature T the incremental forward resistance is a function of k only. Hence, the requirement for balance is that the diode constants k should be equal, which means equal slope of the logarithmic curves. Also, if the two diodes follow the above relation, the incremental resistances will be equal for all values of temperature and so no unbalance can take place due to temperature variations.

In the presence of a large leakage current, I_L, the expression for the incremental resistance is not a simple function of k, T and i only, but is a complicated non-linear function which also includes I_s and I_L. To find two such diodes with equal incremental resistances over the whole operating region is almost impossible, and in practice it is difficult to achieve satisfactory balancing unless both diodes have low leakage currents, i.e. unless they follow a logarithmic law over the whole operating region (Fig. 3, dotted line).

The final modulator circuit is shown in Fig. 5.

The potentiometer P_1 is used to obtain a suitable ratio, a, and C_4 and P_2 are used to equalize the phases and magnitudes of the diode branch impedances during the non-conducting period. With these three controls a balance can quickly be obtained by trial and error. It is important to realize that, not only a.c. balance is required, but also that no direct voltage should exist at the junction point N. This can be checked by

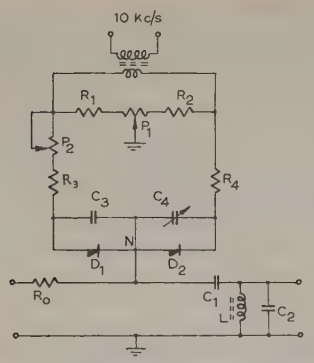


Fig. 5.—Complete modulator circuit.

R_0 680 kΩ.
R_1 R_2 = 100 Ω.
R_3 R_4 = 100 kΩ.
C_1 0·001 μF.
C_2 250 pF.
C_3 20 pF.
C_4 25 pF.
L 0·8 H.
P_1 5 Ω.
P_2 10 kΩ.
D_1, D_2 Silicon diodes type HD 6007.

open-circuiting the input terminal of the modulator temporarily; with perfect balance no output change will be caused.

Variations in the forward resistance of the diodes will be reduced by R_3 and R_4, and the lower the diode resistance the greater this reduction will be; a low forward resistance implies a high value of k in eqn. (2). In spite of this reduction, a simple matching of the diodes may be advisable. The value of R_3 and R_4 should be such as to limit the operation to the logarithmic region of the diode characteristics and yet be several times smaller than R_0. The variable capacitor will balance capacitive differences and, as shown in the next Section, will reduce the effect of variations of diode resistance in the non-conducting direction.

(5) MODULATOR ZERO STABILITY

Since the modulator is essentially a bridge circuit, a change in any of the branches will disturb the balance. If only wire-wound resistors (temperature coefficient 0·002% per deg C) are used and all capacitors are silvered-mica (0·002% per deg C) or air insulated, the drift due to changes in the diode characteristics is far larger than that due to changes in the other components. Also, in bridge circuits the output due to unbalance is proportional to the applied voltage, and so the switching voltage must be as small as possible, and yet sufficiently large to cause conduction in the forward direction.

During the conducting period, drift due to changes in diode resistance is considerably reduced by the large fixed resistors in series with the diodes. The only significant cause of variations of diode characteristics is temperature change and measurements have shown that the drift is less than 0·02 mV/deg C. So far, no ageing effects have been observed.

During the non-conducting period, capacitive as well as resistive components are of importance. Since the balancing capacitor is several times larger than the capacitance of the diode, the total capacitance is effectively constant and only changes in resistance need be considered.

Consider the modulator unbalanced in the non-conducting direction to such an extent that an unbalance voltage, v_{on}, would appear at N in the absence of all capacitive effects. Then with a suddenly applied blocking voltage, V, it was shown in Section 4.1 that

$$v_{on} = V\frac{R_0}{R_0 + R_p}\left(\frac{1}{1+a} - \frac{1}{1+c}\right)$$

Owing to the shunt capacitance, the voltage v_t which actual appears at N is a function of time given by (Section 14)

$$v_t = v_{on}(1 - \varepsilon^{-t/C_pR_T})$$

where C_p is the parallel combination of the capacitance acro the two diodes and $R_T = R_pR_0/(R_p + R_0)$.

The initial rise of voltage per second is then

$$\frac{v_{on}}{C_pR_T} = \frac{V}{C_pR_p}\left(\frac{1}{1+a} - \frac{1}{1+c}\right)$$

which is independent of the input resistance, R_0.

Thus, for a sufficiently high switching frequency, the unbalanc in the non-conducting direction is mainly determined by th time-constant C_pR_p, and can be improved only by increasin this term. C_p is part of the tuned circuit and cannot be increase indefinitely; also, since it is effectively in series with R_3 and R_4 a large balancing capacitor will greatly increase the losses c the circuit. R_p, however, is a property of the particular diode used and should be as high as possible. It is interesting to com pute the unbalance for the modulator in question at the end of blocking period for $R_0 = 1$ megohm and $R_0 = \infty$. With a switching frequency of 10 kc/s, $t = 50$ microsec; also, $R_p = 10^9$ ohms and $C_p = 30$ pF,

$$v_t = V\left(\frac{1}{1+a} - \frac{1}{1+c}\right)0{\cdot}812 \times 10^{-3} \text{ volt for } R_0 = 1 \text{ megohm}$$

$$v_t = V\left(\frac{1}{1+a} - \frac{1}{1+c}\right)1{\cdot}66 \times 10^{-3} \text{ volt for } R_0 = \infty.$$

It is clear that at 10 kc/s the unbalance due to changes in diod resistance in the non-conducting direction is reduced by a factor of about 1000 and is almost independent of the inpu resistor, R_0.

The effects of drift arising in the transit period are not reduced by the modulator circuit as are those which occur during the period of full conduction and complete non-conduction. Experimental results tend to show that this effect can be made small only by using diodes whose characteristics are essentially logarithmic, as mentioned in Section 4.

In conclusion, it may be said that the modulator output due to unbalance is mainly due to changes in diode characteristics with temperature, but fortunately such changes are completely reversible and give rise to no long-term drift.

(6) THE SWITCHING-VOLTAGE SUPPLY

The waveform of the switching voltage is not critical and both square-wave and sine-wave voltages have been tried. There was no significant difference in modulator performance and so for practical reasons, such as transformer design, the sine wave was chosen for normal use.

Since the modulator employs a resonant circuit, the frequency of the switching voltage must be kept reasonably constant. Also, any changes in its magnitude will disturb the balance (a 10% change may cause a drift of up to 0·5 mV), and if a large number of drift correctors are to be supplied from the same generator, a low source impedance is required. In Utac, for which this drift corrector was designed, a 10 kc/s crystal oscillator is available for timing purposes and this is used to supply a simple a.c. regulator as shown in Fig. 6. A portion of the a.c. output is rectified and compared in the feedback path with a constant voltage derived from a reference tube; any difference-voltage is used to control the gain of a pentode and thus maintain the output voltage constant in spite of input and load variations (the regulation at full load is less than 2%). The tuned circuit

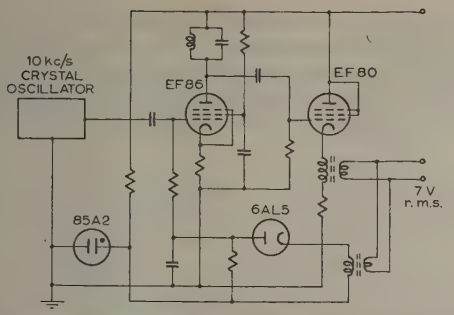


Fig. 6.—10 kc/s regulated source used to supply the switching voltage.

which forms the pentode load ensures a sinusoidal output, free from distortion. The switching power required for each drift corrector is about 10 mW and so a regulator of 1 watt output can supply up to a hundred drift correctors.

Although the use of an a.c. regulator is justified where a number of drift correctors have to be supplied, it should be stressed that the requirements are not extremely critical and a normal standard-frequency oscillator can be used quite successfully.

(7) THE COMPLETE DRIFT CORRECTOR

The modulator described in Section 4 has been used to build drift correctors specifically designed to work in conjunction with the high-gain d.c. amplifier used in Utac. They consist of an input filter, modulator, a.c. amplifier, demodulator and output filter, and the individual sections, shown in Fig. 7, are described below.

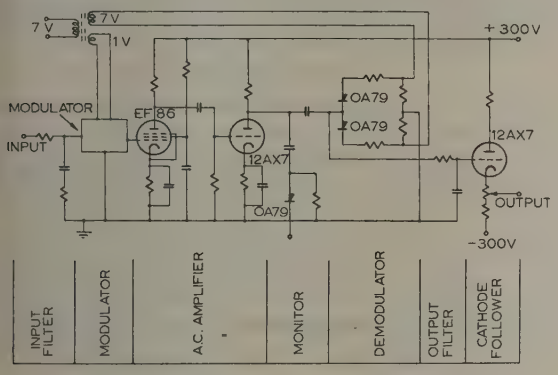


Fig. 7.—Circuit diagram of complete drift-corrector.

The input filter is used to shape the frequency response of the complete drift corrector, and to reduce the carrier-frequency leak back to the input of the amplifier. Its detailed design depends upon the precise application of the drift corrector.

The a.c. amplifier is *RC* coupled, using a pentode as the input stage in order to avoid Miller capacitance effects; this is followed by a triode which feeds directly into the demodulator. The total a.c. gain is 1 600. The only design consideration is that the amplifier should introduce no phase shift at the switching frequency, and for reasons given below, the amplifier gain should remain constant in operation within about 10%.

The demodulator employs two germanium diodes in the same type of bridge circuit as is used in the modulator. No adjustments are required here and normal 10% carbon-composition resistors are satisfactory. The switching voltage is provided from the same transformer which supplies the modulator.

The output filter is a simple *RC* integrating network having a break frequency of 4 c/s.

In Fig. 7 it can be seen that the output filter is directly connected to a cathode follower; this is employed owing to the particular requirements of the d.c. amplifier used, and the output is taken via a potentiometer which controls the output level of the drift-corrected amplifier. The overall d.c. drift-corrector gain is 400, and since the drift of the d.c. amplifier used does not exceed 40 mV, this now corresponds to a maximum drift of only 0·1 mV. This is of the same order as the long-term zero stability of the modulator, and so there is no reason to use a higher drift-corrector gain.

Before demodulation, a small portion of the a.c. signal is rectified and the voltage obtained is a measure of the a.c. amplifier output; this can be displayed on an output meter for monitoring purposes. Since the modulator balance can never be perfect, there is always some residual unbalance voltage at the amplifier input: thus, even when the modulator input is zero, a meter reading will be obtained which indicates this residual voltage. There is an equivalent output from the phase-sensitive demodulator, and this can be balanced out only by using the output potentiometer. However, should the output of the a.c. amplifier change under these circumstances, a form of drift arises; such a change can be caused by drift in the modulator or a change in the gain of the a.c. amplifier. In practice, the residual voltage may be equivalent to up to 1 mV d.c. input, and so a 10% change in amplifier gain can cause a drift of 0·1 mV referred to the input. Normally, a change in modulator balance will cause an increased meter reading, and a reduction in amplifier gain, a reduction in the meter reading; also an overload or equipment failure is immediately apparent.

Once the drift corrector has been set up, the only control which requires adjustment is the output potentiometer; such an adjustment becomes necessary only when a different d.c. amplifier is used or when any of the drift-corrector components have been replaced.

(8) FREQUENCY RESPONSE

As was pointed out in Section 1, the drift corrector is connected so that at low frequencies it is in series with the d.c. amplifier and at higher frequencies the d.c. amplifier alone is used [Fig. 1(*a*)]. The overall gain as a function of frequency is given by

$$KG(j\omega) = [1 + K_2G_2(j\omega)]K_1G_1(j\omega)$$

A logarithmic plot of $K_1G_1(j\omega)$, the frequency-response function of the d.c. amplifier, is given in Fig. 8.

So that the combined amplifier is stable in the presence of feedback, the overall phase lag must be less than 180° at all frequencies at which the combined gain is greater than unity. For this to be so for the given d.c. amplifier, the phase shift due to the term $[1 + K_2G_2(j\omega)]$ should be zero for all frequencies above the lowest break frequency of the d.c. amplifier, which is 400 c/s. This will be the case if the gain of the drift corrector is small compared with unity for all frequencies higher than 400 c/s, and so the design criteria for the frequency response of the drift corrector are:

(*a*) Gain to be less than unity for all frequencies greater than 400 c/s.
(*b*) Phase shift to be less than 180° at all lower frequencies.

The actual frequency response is shown in Fig. 9. The attenuation of 9 dB per octave is obtained by suitably choosing the input and output filters, and the attenuation of 12 dB per octave above 300 c/s is due to the output filter and the resonant effect of the tuned circuit. It is seen from Fig. 9 that the drift-corrector gain becomes unity at 200 c/s, and with the d.c.

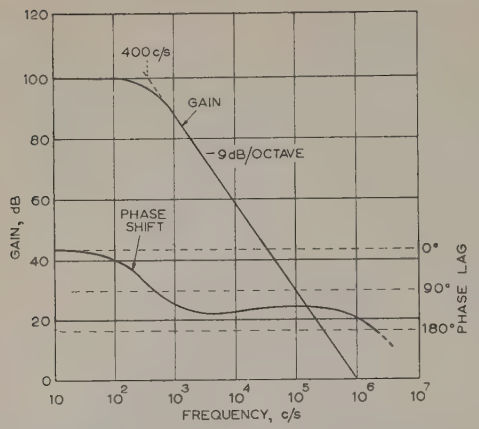


Fig. 8.—Gain and phase response of the d.c. amplifier for which the drift corrector was designed.

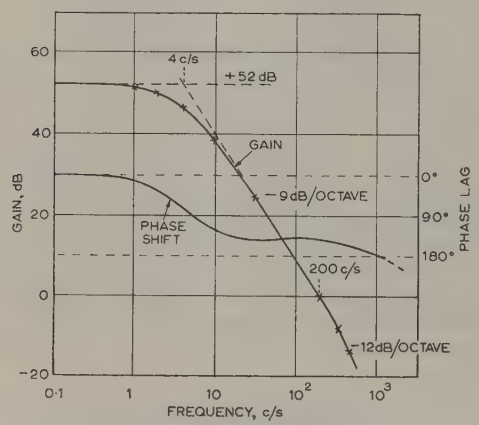


Fig. 9.—Gain and phase response of the drift corrector.

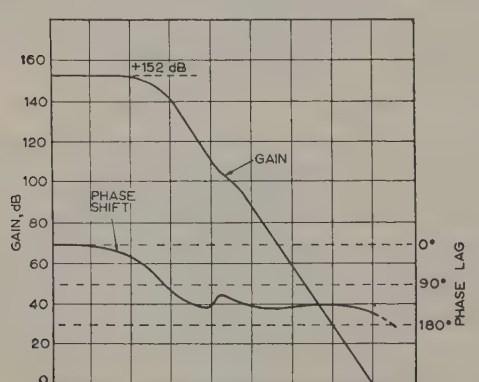


Fig. 10.—Gain and phase response of the complete drift-corrected amplifier.

amplifier this gives the combined frequency response of Fig. 10, which is satisfactory.

(9) DRIFT CORRECTOR IMPERFECTIONS

This Section deals with the effects upon computation of the particular drift corrector described in Section 7.

(9.1) Input Impedance

With the drift-corrected amplifier used in a circuit as show in Fig. 11, for any input voltage, v, an output voltage, v'', w be established; its size will depend upon the computing com ponents Z_i and Z_f. Also, an error voltage, v', will be set up a

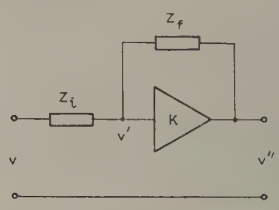


Fig. 11.—Drift-corrected amplifier in a normal computing circuit.

the input of the amplifier. For low frequencies, the paralle branch of the input filter is ineffective, and the error voltag will work into the impedance offered by the modulator plus th series resistor of the input filter only (Figs. 5 and 7). During the non-conducting period the modulator impedance is effec tively the impedance of the tuned circuit at resonance, which is several times larger than the input resistance, R_0. During the conducting period it is R_0 plus the series resistance of the filter.

Thus, a mean current, given very nearly by

$$i' = \tfrac{1}{2}v'/(R_0 + 270\,000) \text{ amp}$$

will flow and so the input resistance may be taken to be $2(R_0 + 270\,000)$, which is approximately 2 megohms.

For all higher frequencies, the parallel branch of the input filter must be considered, and at frequencies well above 10 c/s the input impedance will again become resistive, the approximate value being $(270 + 47)10^3 = 0{\cdot}317$ megohms.

(9.2) Modulator Drift

Consider the drift-corrected amplifier shown in Fig. 11, with the input lead connected to earth; in the absence of drift, $v = v'' = 0$. If a drift voltage appears at the junction N of the modulator (Fig. 5), it will give rise to an output voltage, v'', which in turn will set up an error voltage, v', tending to reduce the effect of the drift. The input impedance offered to v' depends to some extent upon whether the modulator drift is due to unbalance in the conducting or the non-conducting period (see Section 4). In practice, however, this has no significant effect and modulator drift can be considered in exactly the same way as that arising at the input of a normal d.c. amplifier.

(10) PERFORMANCE

During the development of the drift corrector a number of commercial types of silicon diode were investigated in an attempt to find the most suitable one. These tests were certainly not exhaustive, but it was found that certain types contained a reasonable proportion of diodes which had characteristics of the type indicated in Fig. 3 (dotted line). Thus, from a total of 30 silicon diodes of a suitable make, about half were suitable for use in the modulator. The d.c. forward characteristics were measured and the diodes matched in pairs.

A total of ten drift correctors has been built and the performance observed over a long period when used in conjunction with the Utac d.c. amplifier under normal operating conditions.

After operation for over four months, none of the drift correctors showed a drift, referred to the input, larger than

2 mV, the average being 0·1 mV; these values were measured ter the computer had reached a steady operating temperature. he maximum value of drift at first switching on was 1 mV, t this figure could be reduced to less than a half if the heater ower was left on when the rest of the computer was switched f.

It is thought that better performance could be obtained if me form of temperature control were to be employed. The erage temperature coefficient of the drift corrector on its own 0·03 mV/deg C, the temperature being measured within the odulator compartment.

The low-frequency gain of the drift-corrected amplifier is 0×10^6 up to 4 c/s; for higher frequencies the gain falls at 9 dB er octave. The effective input resistance is 2 megohms and the tal noise level, referred to the input, is 1 mV peak-to-peak, ainly at 50 c/s. The switching power required is 10 mW per rift corrector.

(11) CONCLUSIONS

The above results show that silicon diodes can be used uccessfully in drift correctors for computing amplifiers. Tem- erature is the only significant cause of drift, the effect being ompletely reversible. The circuit developed tends to reduce ll drift except that arising during the transit period from full onduction to non-conduction; this can be kept small only by sing diodes with logarithmic characteristics.

In addition, the silicon-diode modulator has a number of ther desirable features: these include good high-frequency esponse, low noise level, low switching power requirement nd an operational life limited only by the electronic valves.

(12) ACKNOWLEDGMENT

The authors wish to thank Professor R. E. Vowels of the chool of Electrical Engineering, The University of New South Wales, for permission to publish this work.

(13) REFERENCES

1) Goldberg, E. A.: 'Stabilization of Wide-Band Direct-Current Amplifiers for Zero and Gain', *RCA Review*, 1950, **11**, p. 296.
2) Wass, C. A. A.: 'Introduction to Electronic Analogue Computers' (Pergamon Press, 1955).
3) Arguimbau, L. B.: 'Vacuum Tube Circuits' (Wiley, New York, 1948).
4) Bubb, F. W.: 'A Network Theorem and Its Application', *Proceedings of The Institute of Radio Engineers*, 1951, **39**, p. 685.
5) Cutler, M., and Bath, H. M.: 'Surface Leakage Current in Silicon Fused Function Diodes', *ibid.*, 1957, **45**, p. 39.
6) Baldwin, E. M., and Mates, J. A.: 'Silicon Junction Diodes', Electronic Components Conference Proceedings, 1955, p. 53.
7) Moody, N. F.: 'A Silicon Junction Diode Modulator', *Electronic Engineering*, 1956, **28**, p. 94.

(14) APPENDIX: UNBALANCE DUE TO CHANGE OF DIODE BACK-RESISTANCE

In Section 4 it was shown that the unbalance due to a change of diode resistance in the non-conducting direction in the absence of all capacitive effects is

$$v_{on} = \frac{VR_0}{R_0 + R_p}\left(\frac{1}{1+a} - \frac{1}{1+c}\right)$$

The circuit capacitance, however, will considerably affect this unbalance. The time-constant of the series resistors R_3 and R_4 (Fig. 4) and the balancing capacitors is only 2×10^{-6} sec, and in the following analysis can be neglected. Thus, the modulator circuit can be simplified as shown in Fig. 12(*a*), where R', R''

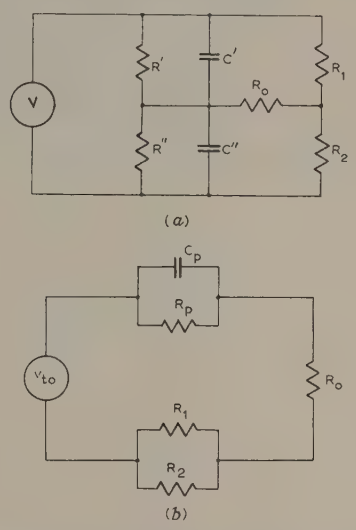


Fig. 12.—Modulator circuits.
(*a*) Simplified circuit applicable only when the diodes are non-conducting.
(*b*) Equivalent circuit for a step input of switching voltage in the non-conducting direction.

and C', C'' are the respective resistances and capacitances of the two diodes. The conditions for balance can be stated as

$$a = \frac{R_1}{R_2} = \frac{R'}{R''} = \frac{C''}{C'}$$

In the following it is assumed that $a = R_1/R_2 = C''/C'$ but $R'/R'' = c \neq a$, and the output, v_t, developed across R_0 is derived for a suddenly applied blocking voltage of magnitude V.

The open-circuit value of v_t is given in Laplace transform notation by

$$v_{to}(p) = V(p)\left[\frac{R_2}{R_1+R_2} - \frac{\dfrac{1}{\dfrac{1}{R''}+pC''}}{\dfrac{1}{\dfrac{1}{R''}+pC''} + \dfrac{1}{\dfrac{1}{R'}+pC'}}\right]$$

$$= V(p)\left[\frac{R_2}{R_1+R_2} - \frac{R''(1+pC'R')}{(R'+R'')(1+pC_pR_p)}\right]$$

where $R_p = R'R''/(R'+R'')$ and $C_p = C' + C''$

For a step input $V(p) = V/p$

$$v_{to}(p) = \frac{V}{p}\left[\frac{R_2}{R_1+R_2} - \frac{R''(1+pC'R')}{(R'+R'')(1+pC_pR_p)}\right]$$

or
$$v_{to}(p) = V\left(\frac{1}{1+a} - \frac{1}{1+c}\right)\left[\frac{1}{p} - \frac{1}{p+\dfrac{1}{C_pR_p}}\right] \quad . \quad (3)$$

The circuit can now be represented as in Fig. 12(*b*). Assuming

that R_0 is much larger than $R_1R_2/(R_1 + R_2)$ the voltage v_t developed across R_0 is given by

$$v_t(p) = v_{to}(p) \frac{R_0}{R_0 + \dfrac{1}{\dfrac{1}{R_p} + pC_p}}$$

$$= v_{to}(p) \frac{R_0(1 + pC_pR_p)}{(R_0 + R_p)\left(1 + pC_p\dfrac{R_0R_p}{R_0 + R_p}\right)}$$

The value of $v_{to}(p)$ given in eqn. (3) can now be substitute in this expression, and allows a value of v_t to be obtained as function of time. Thus:

$$v_t(t) = V\frac{R_0}{R_0 + R_p}\left(\frac{1}{1 + a} - \frac{1}{1 + c}\right)(1 - \varepsilon^{-t/C_pR_T})$$

or $$v_t(t) = v_{on}(1 - \varepsilon^{-t/C_pR_T})$$

where $$R_T = R_0R_p/(R_0 + R_p)$$

21.385.032.213.13: 621.315.59

The Institution of Electrical Engineers
Monograph No. 347 E
Oct. 1959
©

THE CONDUCTIVITY OF OXIDE CATHODES

Part 7. Solid Semiconduction

By G. H. METSON, M.C., D.Sc., Ph.D., M.Sc., B.Sc.(Eng.), Member, and
EDITH MACARTNEY, M.Sc., B.Sc.

(*The paper was first received* 23rd *October*, 1958, *and in revised form* 2nd *July*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in October*, 1959.)

SUMMARY

The properties of the solid conductivity state below 600° K are considered in the present Part. For comparative purposes the conductivity at 420° K is given the title of 'low-temperature reference conductivity σ_{420}', and this function is studied in relation to its environment. It is found that σ_{420} increases in almost linear fashion with increase of conditioning temperature and that it also increases with increase of chemical activity of the core metal. At constant temperature σ_{420} is invariant over a wide range of applied voltage and current, and shows a high order of stability with time. Perhaps the most surprising result to emerge is the observed vulnerability of the solid conductivity to low-pressure oxygen attack, even at a temperature as low as 300° K.

LIST OF PRINCIPAL SYMBOLS

σ_{420} = Low-temperature reference conductivity (measured at 420° K), micromhos.
σ_{300} = Conductivity measured at 300° K, micromhos.
σ_{1020} = Conductivity of the matrix at normal operating temperature (1 020° K), essentially a vacuum-phase conductivity, millimhos.
T_{int} = Temperature of intercept of conductivity phases, deg K.
T_{con} = Conditioning temperature, deg K.
I_{T700} = Low-temperature total emission (measured at 700° K), mA.
R_d = Resistance of the matrix, ohms.
R_0 = Characteristic value of matrix resistance, ohms.

(1) INTRODUCTION

It is proposed in the present Part to examine some of the properties of the solid semiconduction phase of the oxide matrix. Fig. 1 shows a typical characteristic of the variation of log-conductivity with reciprocal of absolute temperature for a standard S-type* assembly, and the solid conductivity phase can be broadly regarded as covering all temperatures below $(T_{int} - 100)$° K. It was shown in Part 4 that below $(T_{int} - 100)$° K effectively the whole of the matrix current is carried by a solid conduction process, that no act of thermionic emission is involved and that heat is uniformly dissipated throughout the matrix. In Part 6 it was further shown that the solid conduction phase has a negligible magnetoresistance, in marked contra-

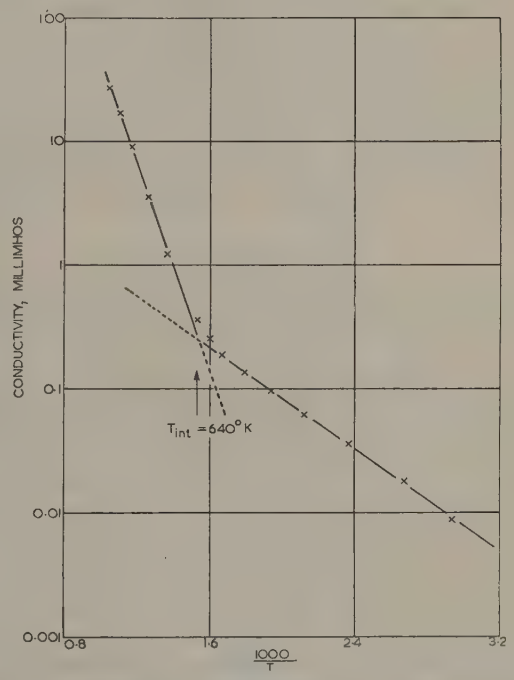


Fig. 1.—Typical temperature characteristic of the conductivity of a platinum-core S-type assembly.

distinction to the vacuum phase. The present Part is concerned with the factors which determine the level of solid conductivity, its relationship to the level of vacuum conductivity and certain of its electrical properties.

(2) SOLID SEMICONDUCTIVITY IN AN S-TYPE ASSEMBLY WITH PLATINUM CORES

(2.1) Thermo-Chemical Action of Platinum on Alkaline-Earth Oxides

Semiconductivity of the oxide matrix results from its activation by an excess of barium metal. This excess may be generated by a direct thermo-chemical core-reducing action of the form

$$X + BaO \rightarrow XO + Ba$$

or by a dissociation of the matrix under the influence of current flow. It is proposed in the present Part to examine only the effects of activation arising from thermo-chemical action.

It is commonly believed that platinum has no useful reducing action at 1 000° K on the alkaline-earth oxides, but experience at Dollis Hill seems to show that this view is incorrect. Con-

* *Specification of standard S-type assembly.*
Cores: Pure platinum, 4% tungsten-nickel or active nickel.
Matrix: Co-precipitated equimolar barium-strontium oxide.
Matrix density: About 1·0.
Matrix thickness: 150μ.
Matrix area: 0·25 cm².
The assemblies are vacuum-processed to the standard schedule detailed in Table 1 of Part 1.

This paper is a continuation of Monographs Nos. 221 R and 243 R, published in February and June, 1957 (see 104 C, pp. 316 and 496), Nos. 268 R, 269 R and 289 R, published in December, 1957, and February, 1958 (see 105 C, pp. 183, 189 and 374) and No. 317 R, published in November, 1958 (see 106 C, p. 55).
Correspondence on Monographs is invited for consideration with a view to publication.
Dr. Metson and Miss Macartney are at the Post Office Research Station.

ventional diodes of the 6D15 type with platinum cores have, for example, run under zero current load for six years without deterioration of total emission.*

It seems probable that losses of excess barium from the matrix by evaporation and random gas poisoning must have occurred during this period and that such losses have been made good by a continuing thermo-chemical action. More direct evidence is afforded by the following experiment. An oxide-coated platinum cathode core fitted with interior heater was suspended by a platinum tape at the centre of an evacuated glass envelope. After the oxide had been heated for 8 hours at 1 700° K a heavy deposit was observed on the cool walls of the glass envelope. The outer fringes of the deposit were a series of colourless layers detectable only by their demonstration of Newtonian rings. These fringe layers were clearly deposits of condensed alkaline-earth oxides. Occupying the centre of the deposit was a heavy grey-black film which disappeared instantly on admitting air to the evacuated envelope. If the alkaline-earth oxide is evaporated in vacuum at 1 700° K from a surface of pure degassed aluminium oxide, the deposit is almost invisible except for Newtonian rings and shows no sign of grey stain. It is assumed that the grey stain in the first sequence is alkaline-earth metal resulting from a thermo-chemical reduction of the oxide by platinum.

(2.2) Thermo-Chemical Activation

Thermo-chemical activation was examined on a standard S-type assembly fitted with platinum cores and processed to the standard schedule set out in Part 1 of the paper. The level of matrix activation, or its concentration of excess barium, is identified with the magnitude of the solid conductivity, which is measured at 420° K, i.e. nearly 300° below the intercept temperature, T_{int}. This conductivity at 420° K will be used extensively and referred to as the 'low-temperature reference conductivity, σ_{420}'. It is a d.c. measurement expressed in micromhos.

The valve was subjected to the following sequence of events. The matrix was set at a conditioning temperature, T_{con}, of 1 050° K for 3 min, allowed to cool to room temperature and then raised to 420° K for measurement of σ_{420}. The whole process was then repeated with the matrix set at a conditioning temperature of 1 100° K for 3 min, allowed to cool and measured for σ_{420}. This sequence was continued until a characteristic of σ_{420}/T_{con} was obtained over a range of T_{con} from 1 050 to 1 600° K. The only current passed through the matrix during the whole sequence was less than the 100 μA necessary to measure the conductivity at 420° K. An example of an experiment carried out on such lines is set out in Fig. 2, which shows that σ_{420} increased with T_{con} to a maximum at about 1 500° K. If the reverse process on a descending temperature scale is now executed, the characteristic follows a new path, as shown in the Figure. This complex behaviour is, in fact, more apparent than real and is due to a one-time sintering action which alters the physical configuration of the matrix reticule. If, for example, the cathode was sintered at the beginning of the experiment by holding the matrix for 3 min at 1 600° K, the characteristic could be cycled indefinitely between 1 050° and 1 550° K in the reversible manner illustrated by Fig. 3. A further sintering occurred when T_{con} exceeded 1 600° K and the characteristic was again displaced. The true form of the σ_{420}/T_{con} characteristic is therefore that given in Fig. 3, which shows that the matrix activation or low-temperature conductivity is a direct function of conditioning temperature.

An alternate experimental approach is illustrated in Fig. 4.

* The type 6D15 diode is described in Part 3 of the paper. It has a 2-watt indirectly-heated cathode and a close-spaced wire-grid anode. The total emission from the cathode at 1 000° K is about 1 amp/cm². This value is estimated from a low-temperature measurement at 700° K, where no electrolytic activation is to be expected.

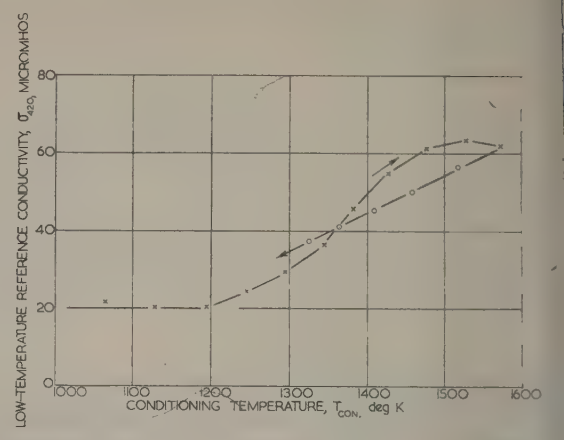


Fig. 2.—Low-temperature reference conductivity as a function of conditioning temperature.

× × × Conditioning temperature increasing.
○ ○ ○ Conditioning temperature decreasing.

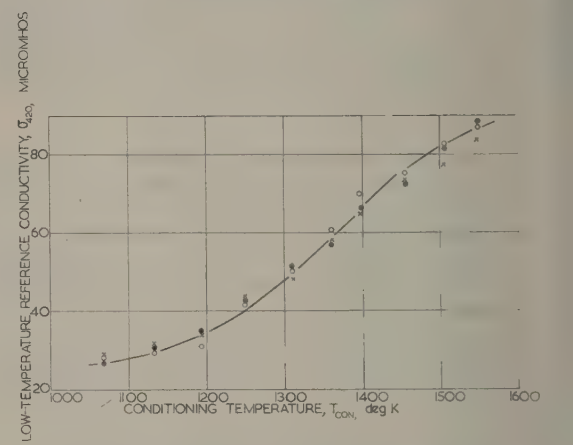


Fig. 3.—Low-temperature reference conductivity of a sintered matrix as a function of conditioning temperature.

× × × First run, temperature decreasing.
○ ○ ○ Second run, temperature increasing.
● ● ● Third run, temperature decreasing.

A standard S-type assembly with platinum cores has been su jected to four different temperature conditionings on a risi temperature scale. After each conditioning the variation of lo conductivity with $1/T$ has been drawn and the resulting fami shows the manner in which an increasing matrix activity affec both solid and vacuum conductivities. The ordinate at 420° gives the characteristic of σ_{420}/T_{con} similar to that of Fig. 3.

(2.3) Reversibility

The mechanism of the activating action is thought to be thermo-chemical one of the form

$$n\mathrm{Pt} + m\mathrm{BaO} \rightarrow \mathrm{Pt}_n\mathrm{O}_m + m\mathrm{Ba}$$

This action at the matrix-core boundary is likely to be irreversib and should lead to irreversibility in the σ_{420}/T_{con} characteristi The fall of solid conductivity with reduction of conditioni temperature observed in Fig. 3 will therefore require explanatio A preliminary experiment to this end is shown in Fig. 5. standard S-type assembly with platinum cores was conditione for 3 min at 1 400° K under zero current load and then measure

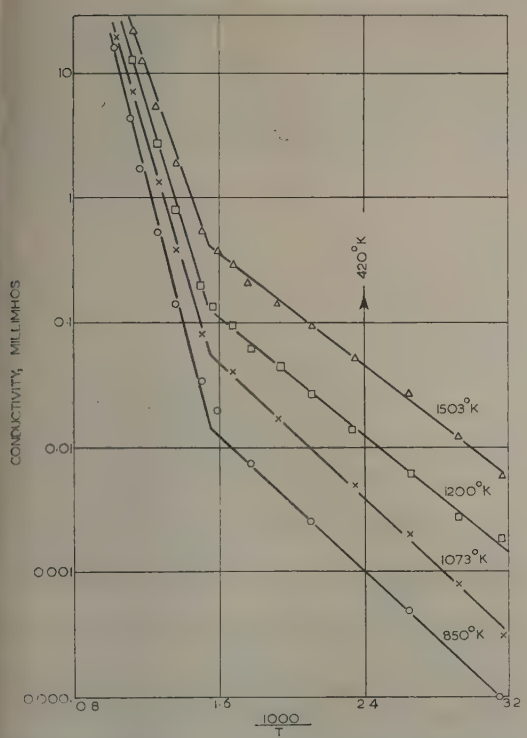


Fig. 4.—Conductivity as a function of reciprocal of temperature after conditioning at various temperatures.

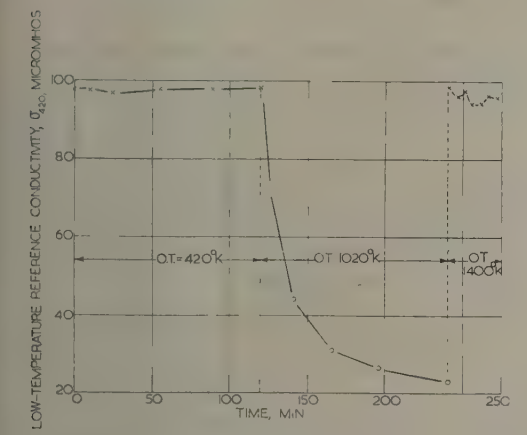


Fig. 5.—Stability of low-temperature reference conductivity as a function of operating temperature—I.

or σ_{420}. The assembly was then set at 420° K for 120 min with ccasional measurement of σ_{420}. The temperature was next ised to 1 020° K and the valve run for a further 120 min with ıtermittent coolings for measurement of σ_{420}. It was finally ın at 1 400° K for 30 min with checks for σ_{420} every 5 min. he characteristics shown in Fig. 5, taken in conjunction with ıose of Fig. 3, bring out the following points:

(*a*) Between 1 020°–1 550° K the solid conductivity, σ_{420}, is a function of conditioning temperature, T_{con}.

(*b*) At 420° K the system is 'frozen' and σ_{420} is stable with time. This has been found to be true for testing times up to 200 hours.

It follows from the above that the system 'freezes' somewhere between 420 and 1 020° K and that σ_{420} is invariant with time below this 'freezing' temperature. The results of a search for the 'freezing' temperature are shown in Fig. 6. A standard

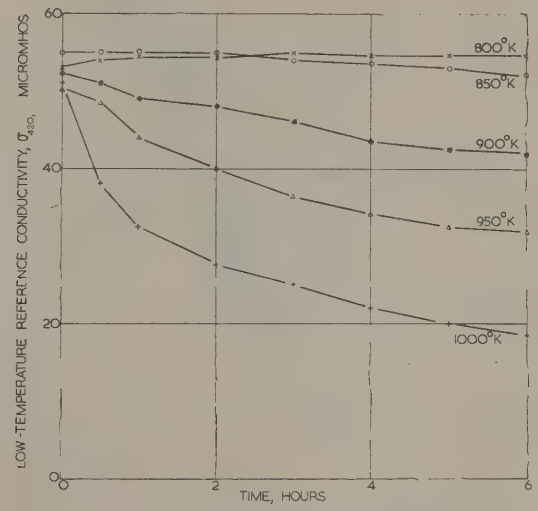


Fig. 6.—Stability of low-temperature reference conductivity as a function of operating temperature—II.

S-type assembly with platinum cores was conditioned for 3 min at 1 400° K and then run for successive periods of 6 hours each under zero current load at temperatures ranging from 400 to 1 000° K in steps of 50°. Measurements of σ_{420} were made at hourly intervals, and the function was found to be invariant with time in the range 400–800° K. To avoid overcrowding the graph these measurements have been omitted from Fig. 6. Instability was just apparent at 850° K, and thereafter became increasingly rapid. After a run in which decay occurs it is, of course, necessary to recondition the tube for 3 min at 1 400° K to regain the common start point. The experiment is a laborious one and was carried out once by each author, who arrived independently at the conclusion that freezing occurs in the range 800–850° K.

The explanation for the freezing effect is thought to be simple and to be in an experiment described in a previous work.[1] In this earlier investigation a conventional form of diode (6D15) with a platinum core was used and the stability of the low-temperature reference total emission, I_{T700}, was examined as a function of the operating temperature, T_{con}. The diodes were processed to a high initial level of activation and run in groups at different operating temperatures with measurement of I_{T700} at suitable intervals. The resulting characteristics of I_{T700}/time at constant T_{con} over a range of T_{con} are set out in Fig. 7 and are directly analogous to the σ_{420}/time characteristics in Fig. 6. Comparison of the two sets of results show that both low-temperature reference conductivity and total emission are stable for all operating temperatures below 800° K, but for higher temperatures both functions become progressively unstable. The cause of the instability of I_{T700} was shown in the previous work[1] to be due to a flow of excess barium metal from the matrix into the platinum core, where it forms a distinctive barium-platinum alloy. This loss of activator by the matrix decreases its low-temperature reference total emission—and likewise its low-temperature reference conductivity. In short, the analogous results of Figs. 6 and 7 are due to activator extraction from the

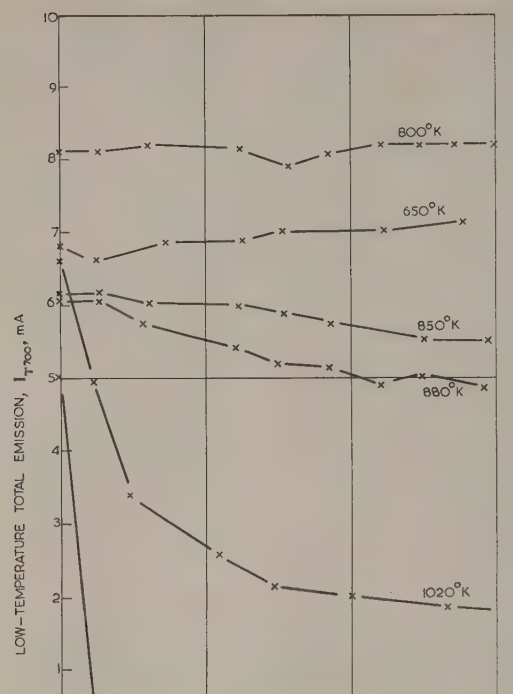


Fig. 7.—Stability of low-temperature total emission of a platinum-core cathode as a function of operating temperature.

matrix by the platinum core metal, and this extraction continues until barium equilibrium exists between matrix and core alloy.

The reversibility of the σ_{420}/T_{con} characteristic in Fig. 3 is now explicable. With increasing temperature the reaction rate between platinum and the alkaline-earth oxides increases, the concentration of excess barium increases, and, in consequence, the low-temperature reference conductivity rises. With decrease of temperature the barium production rate falls and the over-balance of barium left from the previously higher temperature flows into the platinum core to form a barium-platinum interface. On a nickel-alloy core the same sequence is followed on a rising temperature run, but since barium is insoluble in nickel, there is no decrease of σ_{420} on a subsequent falling temperature run, i.e. the overbalance of barium remains in the matrix and the reversible characteristic of platinum shown in Fig. 3 is lost. In short, the platinum system has a barium 'sink' which is absent in the nickel system, and it is this 'sink' which gives to the platinum system a peculiar advantage in demonstrating thermo-chemical activation in a convincing cyclic fashion.

(2.4) Relationship between Magnitudes of Solid and Vacuum Conductivities

The manner in which the vacuum conductivity changes with solid conductivity was easily established, using a standard platinum-core S-type assembly. The relationship required was that between the solid conductivity, σ_{420}, measured at 420° K and the conductivity, σ_{1020}, measured at 1 020 K, which is essentially a vacuum conductivity.

The matrix was first brought into a condition of high activation by a 3 min run at 1 550° K, and then measured for σ_{420} and σ_{1020}. The temperature was then maintained at 1 020° K to induce a fall in activity by barium migration into the core, and the two conductivities were measured as functions of time. A more convenient parameter than σ_{1020} is its reciprocal—matrix resistance R_d—which may be compared with the characteristic resistance R_0. Three typical sets of characteristics are given in Fig. 8.

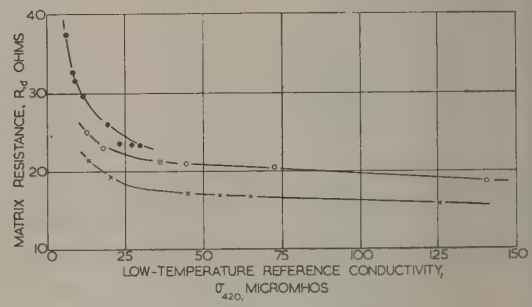


Fig. 8.—Variation of R_d/σ_{420}.

Two points emerge from these characteristics, namely

(*a*) For lower states of activation corresponding to σ_{420} 50 micromhos the vacuum resistance at 1 020° K changes rapidly with change of σ_{420}.

(*b*) For higher states of activation corresponding to σ_{420} 50 microhmos the vacuum resistance is largely unchanged with changing σ_{420}, and tends towards the characteristic resistance.

The reason for the existence of the R_0 state is now apparent the achievement of a quite moderate activation state brings into the range 16 $\pm$ 2 ohms, and any further activation leaves within this range.

(2.5) Action of Oxygen Gas on Solid Conductivity

At 420° K the conductivity of the matrix can be regarded as solid-state phenomenon based on the presence of an excess the alkaline-earth metal. The current may pass through t solid bodies of chains of contiguous particles or may be limit to the surface of such chains. An attempt has been made examine these two simple views by injecting low-pressure oxyg gas into the system. If, at 420° K, excess barium metal is 'froze in the oxide lattice and oxygen atoms are unable to penetra the lattice surface, then gas action should be capable of inte fering only with a surface conductivity.

The valve for the investigation was rather more elaborate th usual, containing a standard S-type assembly with platinu cores, a tungsten spiral coated with barium peroxide as oxygen generator, and a conventional form (6D15) of oxid cathode diode to act as an independent indicator of catho poisoning. The whole arrangement was vacuum processed the usual way, and the two oxide devices were run into a stab state. The temperature of the S-type assembly was then reduc to 420° K and its low-temperature reference conductivity co tinually observed. The indicating diode was meanwhile operat with its oxide cathode set at 1 020° K and its anode curre monitored. When both systems had settled down, a slowl increasing voltage was applied to the peroxide-coated tungst spiral and a simultaneous time record taken of σ_{420} for t S-type assembly and of anode current in the diode. A result set out in Fig. 9, and this can be regarded as typical of a numb of measurements on other valves. Decay of the two chara teristics set in almost simultaneously, and both then fell asympt tically toward zero. The pressure of oxygen producing the

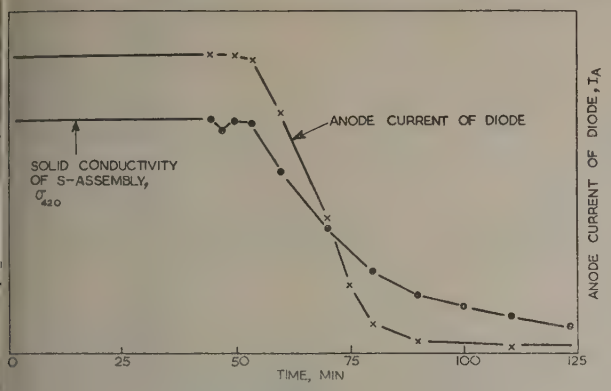


Fig. 9.—Action of oxygen gas on solid conductivity of platinum-core S-type assembly.

effects was not measured, but is estimated from past experience of diode decay rates to be about 10^{-4} mm Hg.

Comment on the several phenomena mentioned in this Section will be reserved for Section 5.

(3) INFLUENCE OF CORE METAL ON SOLID CONDUCTIVITY

(3.1) Comparison of Core Metals

In the previous Section it was suggested that platinum is capable of reducing barium oxide and that a continuous stream of barium is liberated at the oxide-metal boundary. The rate of barium generation must be very small, since careful examination of platinum-cored oxide cathodes which have run for 50 000 hours at 1 020° K show no signs of an evaporated deposit of barium metal on the glass envelope. In contrast, similar valves fitted with the conventional active O-type nickel cores show a very marked deposit of metallic barium* in 3 000 hours of running. The rates of generation in the two cases are thus significantly different.

It is proposed in the present Section to examine the influence of a change in barium generation rate on the magnitude of solid semiconductivity as measured in the standard S-type assembly. To provide a graded rate of generation, the core metals used were: pure platinum, nickel with 4% tungsten, and O-type nickel containing 0·05% magnesium and silicon. The three metals are regarded as feeble, moderate and strong in their respective barium generation rates at a common temperature.

Three sets of valves with the three different core metals were made up to the standard specification and subjected to the standard vacuum processing. Each was then brought into the state of characteristic resistance by thermal means,† and its variation of log-conductivity with reciprocal of absolute temperature was measured. The results are set out in Fig. 10 in the form of four envelopes—a common one for all of the vacuum conductivities and a separate one for each of the three solid conductivities. That all of the valves have roughly the same vacuum conductivity follows, of course, from the concept of the characteristic resistance, but the difference in magnitude of the three solid conductivities is perhaps surprisingly large. In round figures, at 420° K the ratios of the solid conductivities of platinum, nickel-tungsten and O-type nickel are 1 : 10 : 100.

* The film is recognized by its grey colour and its instant disappearance on opening the tube to air. Films of nickel and magnesium are similar in appearance but are unchanged on admission of air.

† The characteristic resistance of an S-type assembly was defined in Part 2 of the paper (Section 1.3).

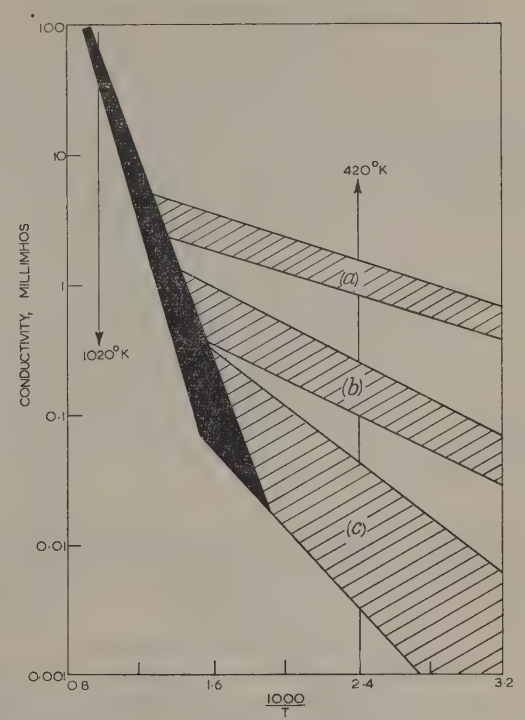


Fig. 10.—Influence of core metal on the magnitude of solid conductivity.

The black area represents the vacuum conductivity phase for all core metals. The shaded areas represent the solid conductivity phase for (*a*) active nickel, (*b*) tungsten nickel, and (*c*) platinum.

(3.2) Action of Oxygen Gas on Solid Conductivity with Active Nickel Cores

The action of low-pressure oxygen gas on the solid conductivity of the S-assembly with O-type nickel cores was examined in the same way as that described in Section 2.5 for the platinum-core case. To avoid the possibility of nickel oxidation, however, the temperature at which the conductivity was observed was reduced from 420° K to the room temperature of about 300° K. The result set out in Fig. 11 is generally similar to the platinum case.

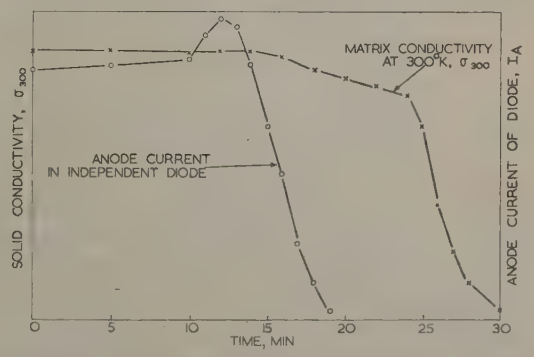


Fig. 11.—Action of oxygen gas on solid conductivity of active nickel-core S-type assembly.

The delay in onset of conductivity decay may possibly be due to a slower rate of penetration of the gas into the S-assembly at the lower testing temperature of 300° K. The pressure of the gas is again likely to be around 10^{-4} mm Hg.

(4) ELECTRICAL PROPERTIES OF THE SOLID CONDUCTIVITY

(4.1) Temperature Coefficients

It will be apparent from Fig. 10 that, apart from different magnitudes of solid conductivity, the three core types have noticeably different temperature coefficients. The magnitude of the coefficient is conveniently expressed as an activation energy in electron-volts and it appears to decrease with increase of chemical activity of core metal. Numerical values are given in Table 1.

Table 1

CHANGE OF LOW-TEMPERATURE ACTIVATION WITH INCREASE OF CHEMICAL ACTIVITY OF CORE METAL

Core metal	Chemical activity	Number of samples	Mean activation energy	Spread
			eV	eV
Platinum ..	Feeble	23	0·23	0·11–0·29
Nickel-tungsten	Moderate	20	0·15	0·09–0·25
O-type nickel	Active	6	0·11	0·07–0·19

The activation energy quoted by Loosjes and Vink[2] for a core metal which appears similar to O-type nickel in composition is 0·09 eV, which compares well with the value of 0·11 eV given in Table 1.

(4.2) Current/Voltage Relationship

The current/voltage relationship of the solid semiconductor has been measured with considerable care at 420° K. To ensure both short- and long-term constancy of temperature, the supplies to the two core heaters were taken from a high-grade d.c. stabilized power-pack.

A first example of a characteristic covering the range 0–300 μA is set out in Fig. 12, where a straight line from the origin has been

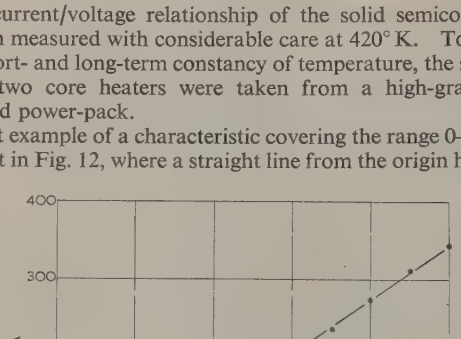


Fig. 12.—Current/voltage characteristic measured at 420° K—I.

drawn through the experimental points. The characteristic is obviously ohmic. A detailed search at 0·1-volt intervals over the range 0–100 μA confirms this ohmic character and shows that, in the absence of thermo-electric potentials due to small difference in core temperature, the characteristic always passes through the origin.

A second example taken over the range 0–1 200 μA is set out in Fig. 13. In this case the maximum power, *VI*, developed in the system at $V = 18{\cdot}0$ volts is 21 mW and this is sufficient to raise the matrix temperature by about 4° K. This temperature increase leads to a rise of conductivity which must result in a departure from linearity of the recorded characteristic. The difficulty of power dissipation can be avoided by using a pulse technique and allowing the system to cool after each application

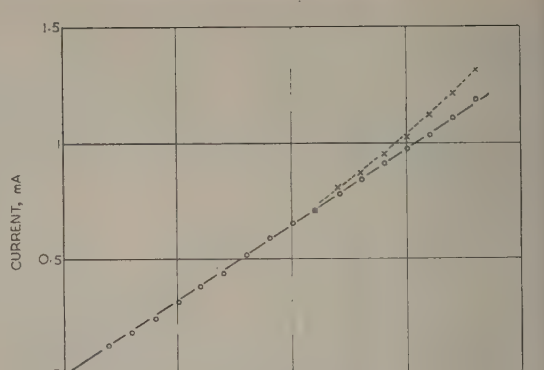


Fig. 13.—Current/voltage characteristic measured at 420° K—II.
—— Measured with short pulses. ---- From continuous measurement.

of voltage. The principal characteristic in Fig. 13 was obtained using a 1 sec voltage application followed by a 1 min cooling period, and the result is linear over the whole current range. The dotted line shows the result of a continuous application of voltage and the departure from linearity becomes apparent for currents greater than 600 μA. The movement from one characteristic to the other is quite slow and the pulse technique therefore easy to apply and apparently trustworthy.

(4.3) Electrical Stability of Solid Conductivity with Time

The electrical stability of σ_{420} with time has been investigated under various current loadings. The measurement requires a high degree of matrix-temperature constancy, and this has been provided. A typical set of results for a platinum-cored S-type assembly is shown in Fig. 14, in which the value of σ_{420} is set

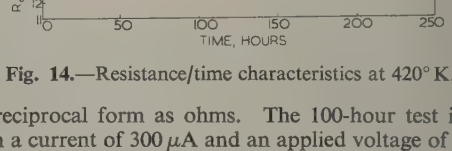


Fig. 14.—Resistance/time characteristics at 420° K.

out in reciprocal form as ohms. The 100-hour test is carried out with a current of 300 μA and an applied voltage of 5·0 volts and the longer-term test at 1 200 μA with a voltage of 18·0 volts. Both results show a high degree of electrical stability of low-temperature conductivity with time. The stability of the system at 18·0 volts is interesting, since this impressed voltage is well above the ionizing potential (about 13 eV) of the common gases.

(4.4) Mechanical Stability

In addition to electrical stability, the solid conductivity at 420° K shows a high degree of mechanical stability. Since the matrix consists of a large number of widely dispersed particles sintered together at point contacts to form a filamentary reticule, the overall mechanical strength is perhaps surprising. An example of a test conducted at 420° K on a standard S-type assembly with nickel-tungsten cores is shown in Fig. 15. A potential of 1 volt is held constant across the matrix and shows an initial value of σ_{420} of 165 micromhos. The test is carried out by striking the glass envelope of the assembly successively in four different quadrantal directions in cyclic fashion and observing the conductivity after each impact. By this means any preferred direction of applied force is sought out and

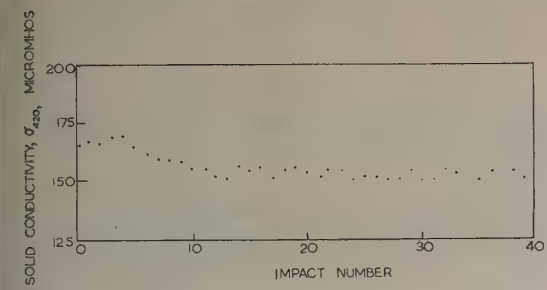


Fig. 15.—Mechanical stability of solid conductivity at 420° K.

examined repeatedly. The impact is applied to the envelope by an ordinary lead pencil swung to give a sharp rap. After a small initial decrease in conductivity the system shows little change over some 30 or more impacts.

(5) SUMMARY AND COMMENT

The experimental observations of the present part may be summarized in the following manner:

(*a*) The low-temperature reference conductivity, σ_{420}, of a platinum-cored S-type assembly increases with increase of conditioning temperature, T_{con}. There appears to be a definite relationship of the form $\sigma_{420} = f(T_{con})$, and equilibrium is achieved rapidly above 1 200° K.

(*b*) For the same operating temperature the function σ_{420} increases with increase of ability of the core metal to reduce the alkaline-earth oxides chemically.

(*c*) There is a definite relationship between the high-temperature vacuum conductivity, σ_{1020}, and σ_{420}. As σ_{420} progressively increases, σ_{1020} makes an asymptotic approach to the reciprocal of the characteristic resistance.

(*d*) The low-temperature conductivity is rapidly and almost wholly destroyed by oxygen. This deactivation has been observed in the temperature range 420–300° K.

(*e*) At 420° K the solid conductivity is observed to follow an accurately ohmic relationship over a wide range of applied voltage.

(*f*) The electrical characteristics of the solid conductivity at 420° K are observed to have a high degree of constancy with time over a wide range of power loading.

(*g*) The solid conductivity at 420° K shows considerable stability under a condition of repeated mechanical shock.

In the work on the platinum-cored S-assembly leading to the relationship $\sigma_{420} = f(T_{con})$ it is suggested that the activation leading to the increase of solid conductivity with increase of temperature is due to a chemical reducing action of platinum metal on barium oxide. This is supported by the evidence of Section 2.1, and is a convenient idea which enables us to include platinum with nickel-tungsten and O-type nickel as purely thermochemical activators. It is, however, equally possible that the purely thermal activations of the platinum-core cases in Figs. 3 and 4 are due to a thermo-dynamical action involving only the matrix wherein the oxide lattice comes into equilibrium with a definite concentration of excess barium for any particular temperature. The experimental work of Moore and Allison,[3] who successfully activated barium-strontium oxide on a chemically neutral base of magnesium oxide lends support to this alternative view.

The second point of comment concerns the rather surprising vulnerability of the solid conductivity at low temperature to low-pressure oxygen attack. It had been anticipated by the authors that the solid conductivity, supposedly dependent on the excess barium atoms uniformly distributed within the crystal lattice of the solid matrix particles, would be largely unaffected by the oxidation of the relatively few excess barium atoms lying on the solid particle surfaces. There is, for example, likely to be only one excess barium atom lying on the surface of a typical matrix particle for every 10 000 contained in the interior. Observation confounds anticipation, however, and, assuming an absence of lattice diffusion at room temperature by both oxygen and barium, it would seem that the solid conductivity is largely a surface phenomenon.

A last point of comment will be made in respect of the form of the $\log\sigma$ versus $1/T$ characteristic in the solid conducting phase. The publications of at least six workers have recorded the characteristic as a linear one and the results of the present Part are in accordance with this view. In a recent publication, however, Higginson[4] suggests that the solid conduction may be more complicated and that the characteristic below 600° K consists of two distinct linear portions with different activation energies. As a result of this suggestion the authors have made a careful recheck of the characteristic in the range of temperature from 600° K down to room temperature, but can find no evidence of any systematic discontinuity. Typical results for a group of active nickel-cored S-type assemblies are set out in Fig. 16,

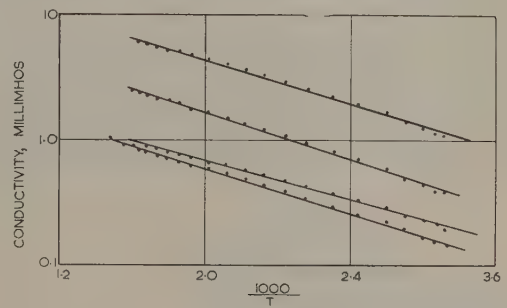


Fig. 16.—Extended measurements of low-temperature characteristics of conductivity.

wherein deviations from linearity are regarded as being due solely to experimental error. One of the authors has on several occasions observed non-linearity of the characteristics in platinum-cored S-type assemblies when the resistance of the system at 1 020° K is well above the characteristic resistance. This phenomenon has been ascribed to the movement of negative ions of gas within the matrix pores, and it disappears when the system is brought into the minimum resistance state by thermal means (i.e. the gas has been expelled). The system used by Higginson is such as the authors would expect to be liable to gas contamination of the matrix. The case for a single-phenomenon view of the solid conductivity is therefore retained by the authors.

(6) ACKNOWLEDGMENTS

Acknowledgment is made to the Engineer-in-Chief of the Post Office for permission to make use of the information contained in the paper. The senior author also wishes to thank Mr. H. Batey for his skilled experimental assistance.

(7) REFERENCES

(1) Metson, G. H.: 'Study of the Long-Term Emission Behaviour of an Oxide-Cathode Valve', *Proceedings I.E.E.*, Paper No. 1790 R, April, 1955 (**102** B, p. 657).

(2) Loosjes, R., and Vink, H. J.: 'Conduction Mechanism in Oxide-Coated Cathodes', *Philips Research Reports*, 1949, **4**, p. 449.

(3) Moore, G. E., and Allison, H. W.: 'Emission of Oxide Cathodes supported on a Ceramic Base', *Journal of Applied Physics*, 1956, **27**, p. 1316.

(4) Higginson, G. S.: 'The Low-Temperature Conductivity of Oxide-Coated Cathodes', *British Journal of Applied Physics*, 1958, **9**, p. 106.

621.314.2.017

The Institution of Electrical Engine
Monograph No. 34
Nov. 19

AN OPTIMUM RATIO OF COPPER LOSSES AND IRON LOSSES FOR A TRANSFORMER WITH VARIABLE LOAD

By G. S. BROSAN, Ph.D., B.Sc.(Eng.), and D. O. BISHOP, Ph.D., B.Sc.(Eng.), Members.

(*The paper was first received* 3*rd April, and in revised form* 18*th June*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in November*, 1959.)

SUMMARY

When a power transformer is loaded so that it is operating at maximum efficiency the copper losses are equal to the iron losses The 'efficiency' is, of course, the 'power efficiency', i.e. the ratio of output power to input power. The same condition applies for maximum 'energy efficiency' only if the load is constant for the period being considered.

With most transformers the load varies considerably over a load cycle and the optimum ratio of copper losses to iron losses calculated on the assumption of constant load for a given period can only be, at least, a rough approximation.

The paper suggests a method of obtaining a more accurate optimum ratio in cases where the load variation of the transformer can be estimated in advance. The assumption is made that the copper resistance is constant.

LIST OF SYMBOLS

p = Load on transformer, per unit.
P_i = Full-load iron loss, per unit.
P_c = Full-load copper loss, per unit.
t = Time.
η = Efficiency.
P_m = Average per-unit load over a given period.
P_R = R.M.S. per-unit load over a given period.
η_m = Maximum efficiency.
η_{mc} = Maximum efficiency based on the assumption of constant load.
P = Constant per-unit load on transformer.
W_0 = Energy output over a given period.

(1) DERIVATION OF OPTIMUM RATIO FOR ANY VARIABLE LOAD

If we consider a transformer delivering a variable per-unit load, p, over a period t_1, the energy output over this period will be

$$W_0 = \int_0^{t_1} p\,dt$$

The energy absorbed in iron losses will be

$$P_i \int_0^{t_1} dt$$

The energy absorbed in copper losses, assuming constant resistance, will be

$$P_c \int_0^{t_1} p^2\,dt$$

Thus the energy efficiency of the transformer over this peri will be

$$\eta = \frac{\int_0^{t_1} p\,dt}{\int_0^{t_1} p\,dt + P_i \int_0^{t_1} dt + P_c \int_0^{t_1} p^2\,dt} \quad . \quad . \quad .$$

For an extreme value, $\frac{\partial \eta}{\partial p} = 0$.

Differentiating with respect to p and equating to zero, have*

$$-P_i + 2P_c \left[\frac{\int_0^{t_1} p\,dt}{\int_0^{t_1} dt}\right]^2 - P_c \left[\frac{\int_0^{t_1} p^2\,dt}{\int_0^{t_1} dt}\right] = 0 \quad . \quad .$$

But $\left[\frac{\int_0^{t_1} p}{\int_0^{t_1} dt}\right]$ is the average load over the given period, i

the average height of the load curve. This is P_m (per uni

$\left[\frac{\int_0^{t_1} p^2\,dt}{\int_0^{t_1} dt}\right]$ is the average height of the curve of p^2 plott

against t over the given period.

Let P_R = Square root of the average height of the curve (p^2, t) over the given per-unit period. This is r.m.s. value of the load over this period (per un

Then

$$\left[\frac{\int_0^{t_1} p^2\,dt}{\int_0^{t_1} dt}\right] = P_R^2$$

Thus eqn. (2) becomes

$$-P_i + 2P_c P_m^2 - P_c P_R^2 = 0 \quad . \quad . \quad . \quad .$$

from which

$$\frac{P_i}{P_c} = 2P_m^2 - P_R^2 \quad . \quad . \quad . \quad . \quad . \quad .$$

Hence, if the variation of p over a load cycle is known or be estimated, P_m and P_R can be found (graphically or otherw and the optimum ratio of iron loss to copper loss can determined.

* EDWARDS, J.: 'A Treatise on the Integral Calculus' (Macmillan, 1930), p. 36

Correspondence on Monographs is invited for consideration with a view to publication.

Dr. Brosan is at Willesden Technical College and Dr. Bishop is at Watford Technical College.

(2) SPECIAL CASE OF CONSTANT LOAD

When the transformer load p is constant

$$P_m = P_R = P$$

Substituting in eqn. (4), we have

$$\frac{P_i}{P_c} = P^2 \text{ or } P_i = P_c P^2 \quad . \quad . \quad . \quad . \quad . \quad (5)$$

But $P_c P^2$ is the copper loss when the per-unit load is P, and thus eqn. (5) gives the commonly used relationship for maximum efficiency, i.e. copper losses equal iron losses.

(3) VALUE OF THE MAXIMUM EFFICIENCY

From eqn. (1)

$$\eta = \frac{\int_0^{t_1} pdt \Big/ \int_0^{t_1} dt}{\frac{\int_0^{t_1} pdt}{\int_0^{t_1} dt} + P_i + P_c \frac{\int_0^{t_1} p^2 dt}{\int_0^{t_1} dt}}$$

or

$$\eta = \frac{P_m}{P_m + P_i + P_c P_R^2} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

For maximum efficiency

$$P_i = (2P_m^2 - P_R^2)P_c$$

Substituting in eqn. (6), the maximum efficiency is

$$\eta_m = \frac{P_m}{P_m + 2P_m^2 P_c - P_R^2 P_c + P_c P_R^2} = \frac{1}{1 + 2P_m P_c} \quad . \quad (7)$$

(4) PRACTICAL CONSIDERATIONS

It is usual, when estimating the most desirable loss ratio, P_i/P_c, to assume that the load, P, on the transformer, is constant over the load cycle and has the value P_m.

In other words, from eqn. (5),

$$\frac{P_i}{P_c} = P_m^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Substituting in eqn. (6), the maximum efficiency when this assumption is made is obtained. Thus

$$\eta_{mc} = \frac{P_m}{P_c P_m^2 + P_m + P_c P_R^2} = \frac{1}{1 + P_c\left(P_m + \frac{P_R^2}{P_m}\right)} = \frac{1}{1 + kP_c} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where

$$k = P_m + \frac{P_R^2}{P_m}$$

It is interesting to compare the maximum efficiency with the more accurate one given by eqn. (7). Thus, from eqn. (7) and (9),

$$\frac{\eta_m}{\eta_{mc}} = \frac{1 + kP_c}{1 + 2P_m P_c} \quad . \quad . \quad . \quad . \quad . \quad (10)$$

and η_m will be greater than η_{mc} if k is greater than $2P_m$.

i.e. if

$$P_m + \frac{P_R^2}{P_m} > 2P_m$$

or

$$P_R > P_m$$

It can be shown that the r.m.s. value of any curve is always equal to or greater than its mean value. Therefore P_R is always equal to or greater than P_m and thus η_m is always equal to or greater than η_{mc}.

In other words, use of the relationship given in eqn. (4) will always result in an improved maximum efficiency when loads are varying.

(5) EXAMPLE OF A SINUSOIDAL LOAD CYCLE

Some indication of the order of the improvement in efficiency is gained by considering a load cycle which varies sinusoidally from 0 to π.

Here

$$P_m = 2/\pi \text{ per unit}$$

and

$$P_R = \frac{1}{\sqrt{2}} \text{ per unit}$$

Thus

$$K = P_m + \frac{P_R^2}{P_m} = 1{\cdot}42$$

Hence, from eqn. (10),

$$\frac{\eta_m}{\eta_{mc}} = \frac{1 + 1{\cdot}42P_c}{1 + 1{\cdot}27P_c}$$

Assuming that the full-load copper loss is 0·03 per unit,

$$\frac{\eta_m}{\eta_{mc}} = 1{\cdot}005$$

There is thus a gain of 0·5% in efficiency by using eqn. (4) to determine P_i/P_c.

(6) CONCLUSIONS

The theory is, of course, approximate because no account is taken of the variation of copper resistance with temperature. It also ignores the relative prices of copper and iron. Nevertheless, if the load cycle of a transformer is known, or even if it can only be estimated approximately, it will always be advantageous to use the ratio of iron loss to copper loss given by eqn. (4) in preference to the more usual one, based on the assumption of constant load.

The saving in energy over a period may be considerable, particularly if the transformer is large and the form factor of the load curve is high.

621.372.22

The Institution of Electrical Engineer
Monograph No. 349
Nov. 195

A NOTE ON THE OPTIMUM DESIGN OF NON-UNIFORM TRANSMISSION LINES

By L. SOLYMAR.

(*The paper was first received 6th December*, 1958, *and in revised form 10th June*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in November*, 1959.)

SUMMARY

The optimum design of non-uniform transmission-lines for a given bandwidth is discussed, and it is stated that the shorter the non-uniform transmission line the more violently its characteristic impedance varies. The concept of complexity of this characteristic impedance function is introduced and a design method is suggested for minimizing the complexity.

Numerical examples are given to illustrate the method, and it is shown that it is possible to design tapers even shorter than the Chebyshev type.

The method is capable of further extension.

LIST OF SYMBOLS

x = Co-ordinate in the direction of propagation.
$V(x)$ = Voltage along the line.
$I(x)$ = Current along the line.
$\rho(x) = \dfrac{V(x)/I(x) - Z_0(x)}{V(x)/I(x) + Z_0(x)}$.
$Z_0(x)$ = Nominal characteristic impedance.
$\gamma = \alpha + j\beta$ = Nominal propagation coefficient.
α = Nominal attenuation coefficient.
β = Nominal phase-change coefficient.
Z_{01}, Z_{02} = Characteristic impedances of the uniform lines connected by the non-uniform line (Fig. 1).
l = Length of the taper.
$F(x) = \frac{1}{2}\dfrac{d}{dx}[\log_\varepsilon Z_0(x)]$.
$u = \beta l, y = 2x/l$ = Variables defined by eqn. (5).
$f(y)$ = A function defined by eqn. (5).
$\rho(-\frac{1}{2}l, u)$ = Reflection coefficient at the point $x = -\frac{1}{2}l$.
Δ = An interval of the u-axis.
δ = Maximum permissible value of the reflection coefficient in the pass band.
B = Complexity of the nominal characteristic-impedance function.
λ_k = Lagrange multiplier.
K, μ_k, C_0, C_1 = Indefinite constants.

(1) INTRODUCTION

One of the principal uses of non-uniform transmission lines is for broadband matching of unequal impedances, e.g. two transmission lines of unequal characteristic impedance. Usually, it is required that the reflection coefficient of the taper shall fall within predetermined limits over a required bandwidth. Recently a number of solutions have been produced aiming at an optimum engineering design. If a taper is required to have a theoretically infinite bandwidth, the Chebyshev taper is undoubtedly the optimum design[1, 2] (i.e. has minimum length); but from an engineering aspect it cannot be the best, since an infinite bandwidth is never demanded. We therefore ask the question, What is the optimum taper design if the bandwidth is finite and specified? The purpose of this study is to investigate this proposition in detail with a view to producing a suitable engineering design.

Without going into the intricacies of the mathematics it can be stated that an optimum taper design cannot exist, for, in an infinite bandwidth, if the taper length is minimized the resulting taper becomes a violently oscillatory function which has no limit; the corresponding physical taper would, of course, not be a practical solution. So far as the engineering aspects of the problem are concerned, it is nevertheless possible to design tapers which are significantly shorter than the Chebyshev taper and yet of simple shape, and the paper indicates the design procedure.

(2) THEORETICAL FUNDAMENTALS

From the familiar transmission-line equations it is found that the function

$$\rho = \frac{V(x)/I(x) - Z_0(x)}{V(x)/I(x) + Z_0(x)}$$

satisfies the first-order non-linear differential equation[3]

$$\frac{d\rho}{dx} - 2\gamma\rho + \frac{1}{2}(1 - \rho^2)\frac{d[\log_\varepsilon Z_0(x)]}{dx} = 0 \quad . \quad . \quad (1$$

If $V(x)/I(x)$ does not differ appreciably from the nominal characteristic impedance, $\rho^2 \ll 1$. This assumption is permissible if the line parameters change slowly and the ratio between the characteristic impedances of the two uniform lines Z_{02}/Z_{01}, is not too large. Introducing the notation

$$F(x) = \frac{1}{2}\frac{d[\log_\varepsilon Z_0(x)]}{dx} \quad . \quad . \quad . \quad . \quad . \quad (2$$

gives a first-order linear differential equation[4] in ρ. The boundary condition (Fig. 1) at $x = \frac{1}{2}l$ is $V(\frac{1}{2}l)/I(\frac{1}{2}l) = Z_{02} = Z_0(\frac{1}{2}l)$

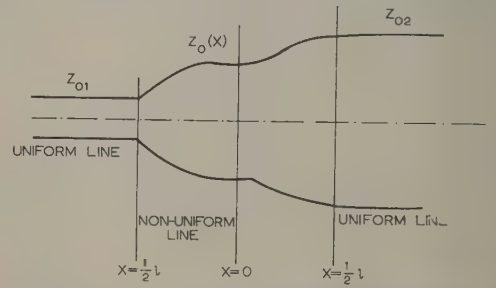


Fig. 1.—Non-uniform transmission line section.

i.e. $\rho = 0$. Under these conditions the differential equation (1) gives

$$\rho(x) = \int_x^{l/2} F(z) \exp\left[-2\int_x^z \gamma(\xi)d\xi\right]dz \quad . \quad . \quad (3)$$

Neglecting losses, $\gamma(x) = j\beta(x)$, a pure imaginary number

Correspondence on Monographs is invited for consideration with a view to publication.
Mr. Solymar is with Standard Telecommunication Laboratories, Ltd.

Further assume that $\beta(x) = \beta$, independent of x but proportional to frequency.

For $x = -\frac{1}{2}l$ eqn. (3) gives

$$\rho(-\tfrac{1}{2}l) = \varepsilon^{-j\beta l}\int_{-\frac{1}{2}l}^{\frac{1}{2}l} F(x)\varepsilon^{-j2\beta z}dx \quad . \quad . \quad . \quad (4)$$

Substituting

$$y = \frac{x}{\frac{1}{2}l};\ \beta l = u;\ F(\tfrac{1}{2}yl) = f(y);\ \rho(-\tfrac{1}{2}l, u) = \Gamma(u)/\varepsilon^{-j\beta l} \quad . \quad (5)$$

gives

$$\Gamma(u) = \tfrac{1}{2}\int_{-1}^{1} f(y)\varepsilon^{-juy}dy \quad . \quad . \quad . \quad . \quad (6)$$

Inverting eqn. (6) (apart from a constant factor) gives

$$f(y) = \int_{-\infty}^{\infty} \Gamma(u)\varepsilon^{juy}du \quad . \quad . \quad . \quad . \quad . \quad (7)$$

The problem is to determine $f(y)$, and hence $Z_0(y)$, for any specified $\Gamma(u)$ function. Obviously the function $\Gamma(u)$ must satisfy the following requirements:

(*a*) Its Fourier transform must vanish outside the $(-1, 1)$ interval.
(*b*) Its absolute value must remain within prescribed limits, $\pm\delta$, in a given interval Δ.
(*c*) At zero frequency its reflection coefficient, $\Gamma(0)$, is determined solely by the characteristic impedances Z_{02} and Z_{01}, irrespective of the nature of the non-uniform transmission line.

Before proceeding with the solution of the integral contained in eqn. (7), subject to the above requirements, it must be stressed that the problem is not simply the determination of the Fourier transform of $\Gamma(u)$. For in the first place $\Gamma(u)$ is not known or given as a function of u, but what is given is that $\Gamma(u)$ must belong to a class of functions satisfying the above requirements.

Similar problems have been encountered in connection with the determination of radiation patterns of certain aerials[5, 6] and the mathematical approach will be very similar; the details of the analysis will therefore be omitted.

In principle the above requirements could be met for an arbitrary (but finite) interval Δ lying arbitrarily near to the $u = 0$ point and for arbitrary small δ. Whether the resulting non-uniform line is a physically realizable structure is another question, and it is safe to say that, other factors being equal, the smaller δ the more wavelike $f(y)$ will be. Moreover, if the $f(y)$ function is rich in high-frequency Fourier components, the approximation (set out at the beginning) of neglecting the higher-order modes is no longer valid. However, judicious use of the method to follow will lead, in most cases, to a satisfactory solution.

From what has been said it will be obvious that an optimum non-uniform transmission line does not exist, since for a given length of line, l, there is an infinity of possible $f(y)$ functions, all of which satisfy the requirements (*a*), (*b*) and (*c*).

(3) OUTLINE OF THE NEW APPROACH

The function attribute called 'complexity' is now introduced and is defined as follows: It is proposed to call the function shown in Fig. 2(*a*) 'simple' and that in Fig. 2(*b*) 'complicated'. It will be observed that large values of $|Z_0''(y)|$ in the interval $(-1, 1)$ are peculiar to a function having large complexity and, in general, vice versa. Consequently, the integral of the square of the second derivative taken over the interval $(-1, 1)$ could

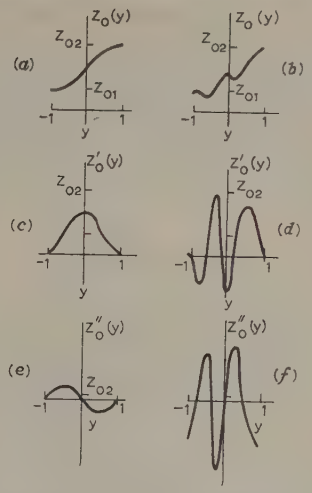


Fig. 2.—Diagrams pertaining to the concept of function complexity.

be taken as a measure of complexity. It will therefore be convenient to define complexity, B, as follows:

$$B = \int_{-1}^{1} [Z_0''(y)]^2 dy \quad . \quad . \quad . \quad . \quad . \quad (8)$$

The physical problem can now be stated as follows. It is required to design a taper [$Z_0(y)$ function] satisfying requirements (*a*), (*b*) and (*c*) and at the same time exhibiting minimum complexity.

Mathematically we seek a function, $Z_0(y)$, which satisfies the conditions (*a*), (*b*) and (*c*) and minimizes the expression (8). From eqn. (2)

$$Z_0(y) = K\exp\left[2\int_{-1}^{y} f(y)dy\right] \quad . \quad . \quad . \quad . \quad (9)$$

where K is an arbitrary constant to be determined.

Since the exponential function is monotonic, a good approximation is obtained by considering only the exponent

$$\int_{-1}^{y} f(y)dy \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The expression to be minimized is therefore

$$B_1 = \int_{-1}^{1}\left\{\left[\int_{-1}^{y} f(y)dy\right]''\right\}^2 dy = \int_{-1}^{1}[f'(y)]^2 dy \quad . \quad (11)$$

For simplicity, assume that $\Gamma(u)$ is a real function, so that $f(y)$ is even. Consequently, the Fourier transforms will contain cosine terms only, and this will simplify further calculations.

The difficult part of the analysis is to satisfy the requirement (*b*), since it is impossible to put it in a form amenable to analysis. However, the difficulty can be circumvented by a number of procedures, as follows.*

In the first instance one could, for example, prescribe the zeros of the $\Gamma(u)$ function at points $u = u_{ak}$ $(k = 1, 2 \ldots n)$, i.e.

$$\int_{0}^{1} f(y)\cos u_{ak}y\,dy = 0\ (k = 1, 2 \ldots n) \quad . \quad . \quad (12)$$

Clearly, if the density of zeros in the interval $(-1, 1)$ is

* It must be realized, however, that the approach need not lead to the best solution, since the auxiliary conditions to be stated are not an exact statement of physical requirements.

increased, there is a strong likelihood for the $\Gamma(u)$ function to remain small in the whole interval.

As an alternative procedure, the values of the function and its derivatives could be prescribed at points $u = u_{bk} (k = 1, 2 \ldots m)$. Let the value of the function at these points be $1/D_{bk}$ of $\Gamma(0)$ (the value at zero frequency), and choose at the same points the derivative of the function to be zero. These conditions can be put in the form

$$\left.\begin{aligned} &\int_0^1 \mathrm{f}(y) \cos u_{bk} y dy = \frac{\Gamma(0)}{D_{bk}} (k = 1, 2 \ldots m) \\ \text{and} \quad &\int_0^1 \mathrm{f}(y) \sin u_{bk} y dy = 0 \end{aligned}\right\} \quad . \quad (13)$$

It would also be possible to assume, at any point u_r, the first q derivatives to be zero. There would then be, in addition,

$$\left.\begin{aligned} &\int_0^1 y^p \mathrm{f}(y) \sin u_r y dy = 0 \quad (p \text{ odd}) \\ \text{and} \quad &\int_0^1 y^p \mathrm{f}(y) \cos u_r y dy = 0 \quad (p \text{ even}) \end{aligned}\right\} \quad . \quad . \quad (14)$$

Evidently the function $Z_0(y)$ can be made to satisfy some or all of these conditions; it would also be made to satisfy any other suitable conditions. Clearly, the problem belongs to the class soluble by the calculus of variations, with the auxiliary conditions in the integral form [e.g. eqn. (12), (13) or (14)].

However, before the method is applied successfully, it is necessary to gain a certain amount of experience and skill. For, if the auxiliary conditions are chosen unskilfully, without any regard to the shape of $Z_0(y)$ curve, the resulting taper, although it will satisfy the bandwidth requirement, will be too complicated in shape (rich in high-frequency Fourier components) to be of any practical use. For any practical case one must therefore seek a compromise between the complexity of the $Z_0(y)$ function and the bandwidth.

To illustrate the method, the auxiliary condition given in eqn. (12) is used. The problem is then formulated as follows:

Required

$$\int_0^1 [\mathrm{f}'(y)]^2 dy = \text{Minimum} \quad . \quad . \quad . \quad . \quad (15)$$

subject to the auxiliary condition

$$\int_0^1 \mathrm{f}(y) \cos u_k y dy = 0 \; (k = 1, 2 \ldots n) \quad . \quad . \quad (16)$$

It can be shown that the necessary condition for the existence of the solution is that the function $\mathrm{f}(y)$ satisfies the Euler–Lagrange differential equation (8). (Whether or not the condition is sufficient will be obvious from physical reasoning.) The appropriate Euler–Lagrange differential equation for this case reads

$$\frac{\partial}{\partial \mathrm{f}(y)} \left\{ [\mathrm{f}'(y)]^2 + \sum_{k=1}^{n} \lambda_k \mathrm{f}(y) \cos u_k y \right\} - \frac{d}{dy} \frac{\partial}{\partial \mathrm{f}'(y)} \left\{ [\mathrm{f}'(y)]^2 + \sum_{k=1}^{n} \lambda_k \mathrm{f}(y) \cos u_k y \right\} = 0 \quad . \quad (17)$$

where the λ_k's are Lagrangian multipliers.

This reduces to

$$\mathrm{f}''(y) = \frac{1}{2} \sum_{k=1}^{n} \lambda_k \cos u_k y \quad . \quad . \quad . \quad . \quad (18)$$

Integrating and introducing the new variables μ_k instead $\frac{1}{2}\lambda_k/u_k^2$ gives

$$\mathrm{f}(y) = \sum_{k=1}^{n} \mu_k \cos u_k y + C_1 y + C_0 \quad . \quad . \quad (1$$

The unknown μ_k are determined with the help of eqn. (1(while the constants C_0 and C_1 are determined by the followi procedure.

The application of the boundary conditions to the nomin characteristic-impedance function yields

$$\left.\begin{aligned} &Z_0(y) = Z_{01} \text{ if } y = -1 \\ &Z_0(y) = Z_{02} \text{ if } y = 1 \end{aligned}\right\} \quad . \quad . \quad (2$$

Substituting in eqn. (9) with $K = Z_{01}$ gives

$$\frac{1}{2} \int_{-1}^{1} \mathrm{f}(y) dy = \frac{1}{2} \log_e \frac{Z_{02}}{Z_{01}} \quad . \quad . \quad . \quad . \quad (2$$

which is a good approximation to the zero-frequency reflecti coefficient provided that Z_{02}/Z_{01} is less than 2 (see Reference 2

The integration constant C_1 can be determined in a number ways, and the logical procedure is to determine it by minimizi the expression

$$\mathrm{B}_1(C_1) = \int_{-1}^{1} [\mathrm{f}'(y)]^2 dy \quad . \quad . \quad . \quad . \quad (2$$

In the particular application, $C_1 = 0$, for simplicity.

(4) AN APPLICATION OF THE METHOD

To be able to compare the numerical results with tho obtained by different methods and already published elsewher it will be convenient to choose $\Gamma(0) = 0{\cdot}2$.

Let the first zero of the $\Gamma(u)$ function be at π, as in the ca of the exponential taper. If the second zero occurs at $3\pi/2$ a the method described above is applied (details of the calculatio can be found in the Appendix), we obtain a curve as shown Fig. 3. It will be observed that in the interval 2·82–5·12 t

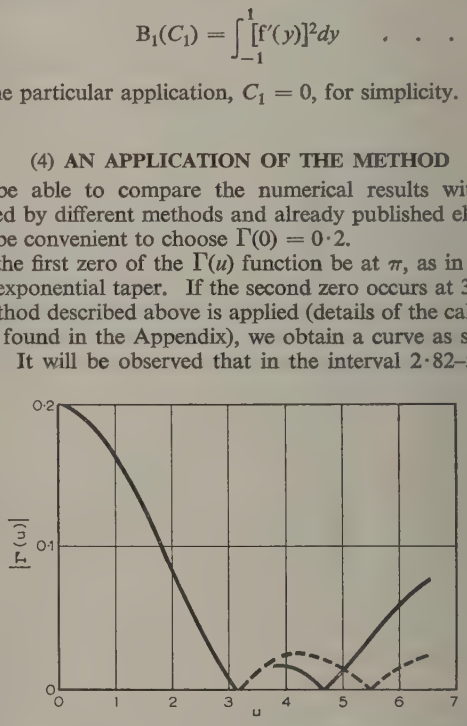


Fig. 3.—Absolute value of the reflection coefficient $|\Gamma(u)|$ as t function of u.

——— Present design. — — — Chebyshev taper.

reflection coefficient is already less than it would have been f the Chebyshev taper having the same first zero.

If the calculations are continued for various values of t position of the first zero, it is found that, in general, the small the first zero the larger will be the value of δ and the mo complicated the $Z_0(y)$ function . Conversely large values of t first zero are associated with a long taper, small bandwidt small values of δ and a simple $Z_0(u)$ function.

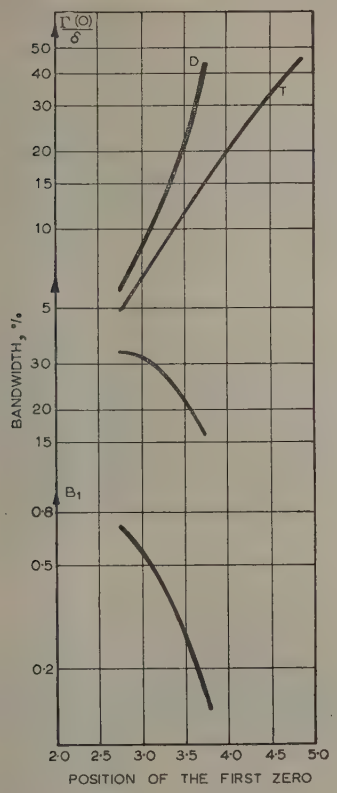


Fig. 4.—Complexity, bandwidth, and $\Gamma(0)/\delta$ for the taper (D) as a function of the position of the first root of the function; for comparison, $\Gamma(0)/\delta$ for the Chebyshev taper (T) is also plotted.

Fig. 4 shows the magnitude of $\Gamma(0)/\delta$, complexity and the bandwidth as a function of the position of the first zero. The results are contrasted with the Chebyschev taper, which, of course, has infinite bandwidth. It transpires from this Figure that for a given value of δ the new taper is substantially shorter than the Chebyshev taper. It can be said that, with the method outlined, a reduction in length of some 10–30% is easily achievable. It must be pointed out, however, that this taper has a finite bandwidth of 20–30%. Nevertheless, this is quite adequate for most practical applications, but, if desired, there is no reason why the computations should not be carried further and an even better result obtained. On the other hand, although mathematically the Chebyshev taper has an infinite bandwidth, in practice this is not so for a number of reasons, e.g. the junction discontinuities have been neglected and the higher-order mode effects disregarded.

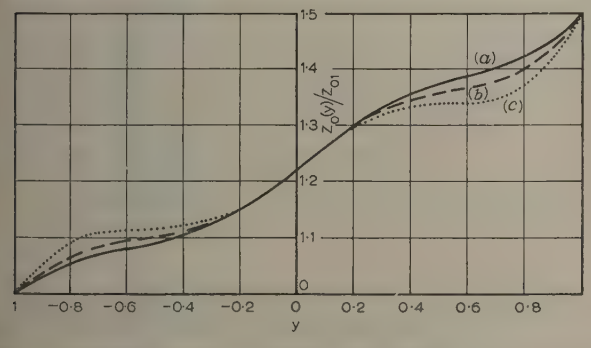


Fig. 5.—Variation of the normalized taper impedance, $Z_0(y)/Z_{01}$, as a function of y, the normalized taper co-ordinate.

(*a*) for $\varepsilon = 0{\cdot}5$.
(*b*) for $\varepsilon = 0$.
(*c*) for $\varepsilon = -0{\cdot}4$.

The resultant $Z_0(y)/Z_{01}$ function of the new taper is shown in Fig. 5 for three different taper lengths [positions of the first zero of the $\Gamma(u)$ function].

(5) CONCLUSIONS

A method based on the variational calculus has been developed and applied to the design of transmission-line tapers. The validity of the method is subject to the following approximations (in common with all other known methods): the effect of higher-order modes is neglected; the ratio of voltage to current in the line does not differ appreciably from the nominal characteristic impedance (this implies that $\rho \ll 1$); the nominal phase-change coefficient, β, is a constant which is the characteristic of the line.

In applying the method to practical cases, two quantities must be given, namely the bandwidth and the maximum permissible value of the reflection coefficient, δ, within it.

The tapers designed by this method have finite bandwidth, but this, as explained above, is no limitation of the method.

In principle, the maximum permissible value of the reflection coefficient, δ, in a given bandwidth could be made as small as desired, but the resulting taper would be very complicated and then the assumption about neglecting the higher-order modes would be invalidated, and in any case the taper would be too complicated to be of any use.

It is also concluded that an optimum design does not exist.

(6) ACKNOWLEDGMENTS

The author wishes to thank Mr. L. Lewin and Dr. A. E. Karbowiak for reading the manuscript.

Acknowledgment is also made to Standard Telecommunication Laboratories, Ltd., for permission to publish the paper.

(7) REFERENCES

(1) KLOPFENSTEIN, R. W.: 'A Transmission-Line Taper of Improved Design', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 31.

(2) COLLINS, R. E.: 'The Optimum Tapered Transmission Line Matching Section', *ibid.*, p. 359.

(3) WALKER, L. R., and WAX, N.: 'Non-Uniform Transmission Lines and Reflection Coefficients', *Journal of Applied Physics*, 1946, **17**, p. 1043.

(4) BOLINDER, F.: 'Fourier Transforms in the Theory of Inhomogeneous Transmission Lines', *Proceedings of the Institute of Radio Engineers*, 1950, **38**, p. 1354.

(5) WOODWARD, P. M., and LAWSON, J. D.: 'The Theoretical Precision with which an Arbitrary Radiation Pattern may be obtained from a Source of Finite Size', *Journal I.E.E.*, 1948, **95**, Part III, p. 363.

(6) KOVACS, R., and SOLYMAR, L.: 'Theory of Aperture Aerials based on the Properties of Entire Functions of the Exponential Type', *Acta Physica Academiae Scientiarum Hungaricae*, 1956, **6**, p. 161.

(7) TAYLOR, T. T.: 'Design of Line-Source Antennas for Narrow Beamwidth and Low Side Lobes', *Transactions of the Institute of Radio Engineers*, 1955, AP–3, p. 16.

(8) WEINSTOCK, R.: 'Calculus of Variations' (McGraw-Hill, New York, 1952).

(8) APPENDIX

The reflection coefficient at $u = 0$ is $\Gamma(0) = 0{\cdot}2$, the prescribed zeros of the $f(y)$ function are $\pi + \varepsilon$ and $3\pi/2$. According to eqn. (19), $f(y)$ can be written in the following form:

$$f(y) = \mu_1 \cos(\pi + \varepsilon)y + \mu_2 \cos\frac{3\pi}{2}y + C_0 \quad . \quad (23)$$

For the determination of μ_1, μ_2 and C_0 we can use eqns. (16) and (21). Hence

$$\left.\begin{aligned} b_{11}\mu_1 + b_{12}\mu_2 + b_{13}C_0 &= 0 \\ b_{21}\mu_1 + b_{22}\mu_2 + b_{23}C_0 &= 0 \\ b_{31}\mu_1 + b_{32}\mu_2 + b_{33}C_0 &= \Gamma(0) \end{aligned}\right\} \quad . \quad . \quad (24)$$

where

$$b_{11} = \int_0^1 \cos^2(\pi + \varepsilon)y\,dy = \frac{1}{2}\left[1 + \frac{\sin 2(\pi + \varepsilon)}{2(\pi + \varepsilon)}\right] \quad . \quad (25)$$

$$b_{12} = b_{21} = \int_0^1 \cos(\pi + \varepsilon)y \cos\frac{3\pi}{2}y\,dy = -\frac{3\pi}{2}\frac{\cos\varepsilon}{(\pi + \varepsilon)^2 - \left(\frac{3\pi}{2}\right)^2} \quad . \quad (26)$$

$$b_{13} = b_{31} = \int_0^1 \cos(\pi + \varepsilon)y\,dy = -\frac{\sin\varepsilon}{\pi + \varepsilon} \quad . \quad (27)$$

$$b_{22} = \int_0^1 \cos^2\frac{3\pi}{2}y\,dy = \frac{1}{2} \quad . \quad . \quad . \quad . \quad (28)$$

$$b_{23} = b_{32} = \int_0^1 \cos\frac{3\pi}{2}y\,dy = -\frac{2}{3\pi} \quad . \quad . \quad . \quad (29)$$

$$b_{33} = \int_0^1 dy = 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad (30)$$

When μ_1, μ_2 and C_0 have been determined from eqn. (24) the reflection coefficient can be calculated with the aid of eqn. (6), whence

$$\Gamma(u) = \mu_1 g_1(u) + \mu_2 g_2(u) + C_0 g_3(u) \quad . \quad . \quad (31)$$

where

$$g_1(u) = \int_0^1 \cos(\pi + \varepsilon)y \cos uy\,dy = \frac{(\pi + \varepsilon)\sin\varepsilon\cos u - u\sin u\cos\varepsilon}{u^2 - (\pi + \varepsilon)^2} \quad . \quad (32)$$

$$g_2(u) = \int_0^1 \cos\frac{3\pi}{2}y \cos uy\,dy = \frac{3\pi}{2}\frac{\cos u}{u^2 - \left(\frac{3\pi}{2}\right)^2} \quad . \quad (33)$$

$$g_3(u) = \int_0^1 \cos uy\,dy = \frac{\sin u}{u} \quad . \quad . \quad . \quad . \quad (34)$$

For different values of ε the $\Gamma(u)$ curve can now be plotted from it the $\Gamma(0)/\delta$ ratio and the bandwidth can be determined. The $Z_0(y)$ curve is calculated from eqn. (9), while the complexity is obtained from eqn. (11).

512.9:621.372.6

The Institution of Electrical Engineers
Monograph No. 350 M
Dec. 1959
©

SOME TESTS FOR THE NUMBER OF POSITIVE ZEROS AND FOR THE NUMBERS OF REAL AND COMPLEX ZEROS OF A REAL POLYNOMIAL

By. O. P. D. CUTTERIDGE, M.Sc.(Eng.), Ph.D., Associate Member.

(*The paper was first received 6th March, and in revised form 11th August*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in December*, 1959.)

SUMMARY

The numbers of positive, real and complex zeros of a polynomial with real coefficients are shown to be related to the number of changes in sign along one or two sequences of determinants whose elements consist of the coefficients of the given polynomial and of its derivative. This in turn is related to the number of negative signs occurring in a certain continued fraction. Some numerical examples illustrate the application of both methods, and the simplification of and the relationship between the two sequences of determinants is also discussed.

(1) INTRODUCTION

In a number of engineering applications it is of importance to be able to investigate the character of the zeros of a polynomial directly from its coefficients, i.e. to determine the number of complex zeros, the number of real zeros and the number of positive zeros that the polynomial possesses without actually calculating them all first.

For example, the roots of an algebraic equation might correspond to the optimum value for a component in an electrical circuit and it is desired to know how many positive, or real, solutions are possible or whether, in fact, any such exist at all. Again, in the synthesis of linear passive driving-point impedances the real part of the impedance must be non-negative for all real frequencies, and checking whether or not this requirement is satisfied for a given function reduces to a consideration of the positive zeros of a polynomial.

The usual mathematical method of investigating these problems is to use Sturm's theorem, details of which are given in most books on the theory of equations.[1, 2] However, Sturm's theorem results in unnecessarily complicated, and hence lengthy, calculations if the numbers of positive and complex zeros are all that is required, and simpler methods are given in the following Section. It is shown that the problem reduces to a consideration of the changes in sign along one or two sequences of determinants whose elements consist of the coefficients of the given polynomial and of its derivative.* This in turn is shown to be equivalent to investigating the number of negative signs occurring in a certain continued fraction. This last result, stated in Section 2 in the form of two theorems on continued fractions, differs from the J-type continued-fraction test given by Wall.[5]

Some numerical examples to illustrate the methods are given in Section 3, and the simplification of and the relationship between the two sequences of determinants is discussed in the Appendix.

* These determinants are the same as those that have been used, for example, by Fuller,[3, 4] to express the conditions for a polynomial to have all its zeros real and negative, and also for a polynomial to have all its zeros real.

Correspondence on Monographs is invited for consideration with a view to publication.

Dr. Cutteridge is at the Manchester College of Science and Technology (Faculty of Technology, University of Manchester).

(2) TWO THEOREMS ON CONTINUED FRACTIONS

Let $$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \quad . \quad . \quad (1)$$

where $a_n, a_{n-1}, \ldots a_0$ are real,

and $$f'(x) = b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \ldots + b_0 \quad . \quad . \quad (2)$$

Then, in general, we can write

$$\frac{f'(x)}{f(x)} = \cfrac{1}{\alpha_1 x + \cfrac{1}{\beta_1 + \cfrac{1}{\alpha_2 x + \cfrac{1}{\beta_2 + \cfrac{\ddots}{\cfrac{1}{\alpha_k x + \cfrac{1}{\beta_k}}}}}}} \quad . \quad (3)$$

where α_i, $\beta_i \neq 0$ and $k = n$ in the absence of multiple zeros of $f(x)$.

The following theorems will now be proved.

Theorem 1.

The number of positive zeros of $f(x)$ is equal to the number of negative β_i minus the number of negative α_i.

Theorem 2.

The number of real zeros of $f(x)$ is equal to $(k - 2N)$ and the number of pairs of conjugate complex zeros is equal to N, where N is the number of negative α_i.

Both Theorems 1 and 2 yield the numbers of distinct zeros in each case, the multiplicity of zeros not being counted.

A proof of these theorems can be obtained from Sturm's theorem and most easily from one of the modified sets of Sturm function given by Trudi;[6] details of these functions are also to be found in Muir.[7]

The functions in question are the set $X_0, X_1, X_2, \ldots X_k$, where

$$\left.\begin{aligned} X_0 &= 1 \quad X_1 = \begin{vmatrix} a_n & a_{n-1} & 0 \\ 0 & b_{n-1} & 1 \\ b_{n-1} & b_{n-2} & x \end{vmatrix} \\ X_2 &= \begin{vmatrix} a_n & a_{n-1} & a_{n-2} & a_{n-3} & 0 \\ 0 & a_n & a_{n-1} & a_{n-2} & 0 \\ 0 & 0 & b_{n-1} & b_{n-2} & 1 \\ 0 & b_{n-1} & b_{n-2} & b_{n-3} & x \\ b_{n-1} & b_{n-2} & b_{n-3} & b_{n-4} & x^2 \end{vmatrix} \\ &\text{etc.} \end{aligned}\right\} \quad . \quad . \quad (4)$$

and
$$\left.\begin{array}{l} b_{n-1} = na_n \\ b_{n-2} = (n-1)a_{n-1} \\ \quad . \quad . \quad . \\ \quad . \quad . \quad . \\ b_0 = a_1 \end{array}\right\} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

At $x = 0$, the functions X_i become

$$\left.\begin{array}{l} X_0(0) = 1 \quad X_1(0) = -\begin{vmatrix} a_n & a_{n-1} \\ b_{n-1} & b_{n-2} \end{vmatrix} \\ X_2(0) = \begin{vmatrix} a_n & a_{n-1} & a_{n-2} & a_{n-3} \\ 0 & a_n & a_{n-1} & a_{n-2} \\ 0 & b_{n-1} & b_{n-2} & b_{n-3} \\ b_{n-1} & b_{n-2} & b_{n-3} & b_{n-4} \end{vmatrix} \\ \text{etc.} \end{array}\right\} \quad . \quad . \quad . \quad (6)$$

Every term X_{2l+1} has a negative sign which can be removed by rewriting the bigradients in eqns. (6) in a form analogous to the Hurwitz determinants. Thus

$$X_0(0) = 1 \; X_1(0) = \Delta_2 \; X_2(0) = \Delta_4 \ldots X_k(0) = \Delta_{2k} \quad . \quad (7)$$

where Δ_i is the ith-order determinant

$$\begin{vmatrix} b_{n-1} & b_{n-2} & b_{n-3} & b_{n-4} & . & . & . \\ a_n & a_{n-1} & a_{n-2} & a_{n-3} & . & . & . \\ 0 & b_{n-1} & b_{n-2} & b_{n-3} & . & . & . \\ 0 & a_n & a_{n-1} & a_{n-2} & . & . & . \\ 0 & 0 & b_{n-1} & b_{n-2} & . & . & . \\ 0 & 0 & a_n & a_{n-1} & . & . & . \\ 0 & 0 & 0 & b_{n-1} & . & . & . \\ & . & . & . & & & \end{vmatrix} \quad . \quad . \quad (8)$$

The number of changes of sign in the sequence

$$1, \Delta_2, \Delta_4, \Delta_6, \ldots \Delta_{2k} \quad . \quad . \quad . \quad . \quad . \quad (9)$$

is equal to the number of negative terms in the sequence

$$\Delta_2, \Delta_2\Delta_4, \Delta_4\Delta_6, \ldots \Delta_{2k-2}\Delta_{2k} \quad . \quad . \quad . \quad (10)$$

Using the same notation, the highest powers of x and corresponding coefficients in the successive functions X_i are

$$1, a_n b_{n-1} x, a_n \Delta_3 x^2, a_n \Delta_5 x^3, a_n \Delta_7 x^4, \ldots a_n \Delta_{2k-1} x^k \quad . \quad (11)$$

Since a_n and b_{n-1} have the same sign the number of changes of sign in sequence (11) when $x = +\infty$ is equal to the number of negative terms in the sequence

$$a_n^2 b_{n-1}\Delta_3, a_n^2\Delta_3\Delta_5, a_n^2\Delta_5\Delta_7, \ldots a_n^2\Delta_{2k-3}\Delta_{2k-1} \quad . \quad (12)$$

which in turn is equal to the number of negative terms in the sequence

$$b_{n-1}\Delta_3, \Delta_3\Delta_5, \Delta_5\Delta_7, \ldots \Delta_{2k-3}\Delta_{2k-1} \quad . \quad . \quad (13)$$

The quantities α_i and β_i in the continued fraction in eqn. (3) have been given by Bader[8] as

$$\left.\begin{array}{ll} \alpha_1 = \dfrac{a_n}{b_{n-1}} & \beta_1 = \dfrac{\Delta_1^2}{\Delta_2} \\ \alpha_2 = \dfrac{\Delta_2^2}{\Delta_1\Delta_3} & \beta_2 = \dfrac{\Delta_3^2}{\Delta_2\Delta_4} \\ \alpha_i = \dfrac{\Delta_{2i-2}^2}{\Delta_{2i-3}\Delta_{2i-1}} & \beta_i = \dfrac{\Delta_{2i-1}^2}{\Delta_{2i-2}\Delta_{2i}} \\ i = 2, 3, \ldots k & \end{array}\right\} \quad . \quad . \quad (14)$$

Thus the number of negative terms in sequence (13) is equal to the number of negative α terms, and the number of negative terms in sequence (10) is equal to the number of negative β terms. But as the functions X_i are a set of Sturm functions the number of sign changes in sequence (9) minus the number of sign changes, when x is positive, in sequence (11) is equal to the number of positive zeros of $f(x)$. Hence the number of positive zeros of $f(x)$ is equal to the number of negative β terms minus the number of negative α terms, as stated in Theorem 1.

Consider now the total number of changes in sign in two sequences A and B, sequence A being an arbitrary sequence of $(k+1)$ positive and negative terms and sequence B obtained from sequence A by changing the signs of the 2nd, 4th, 6th, . . . terms. Evidently the sum of the number of sign changes in sequences A and B must be k. Thus if N is the number of sign changes in sequence (11) when $x = +\infty$ then $(k - N)$ must be the corresponding number of sign changes when $x = -\infty$, and so the number of real zeros of $f(x)$ must be $(k - 2N)$. As k is the total number of distinct zeros, $f(x)$ must have N pairs of conjugate complex zeros.

But as the number of sign changes, when x is positive, in sequence (11) is equal to the number of negative α terms the number of real zeros of $f(x)$ is equal to $(k - 2N)$ and the number of pairs of conjugate complex zeros is equal to N, where N is the number of negative α terms, as stated in Theorem 2.

(3) NUMERICAL EXAMPLES

(3.1) Example 1

Take
$$f(x) = x^3 + 2x^2 + 3x + 2 \quad . \quad . \quad . \quad . \quad (15)$$

then
$$f'(x) = 3x^2 + 4x + 3 \quad . \quad . \quad . \quad . \quad (16)$$

and
$$\frac{f'(x)}{f(x)} = \cfrac{1}{\cfrac{x}{3} + \cfrac{1}{\cfrac{9}{2} + \cfrac{1}{-\cfrac{2}{15}x + \cfrac{1}{-\cfrac{25}{6} + \cfrac{1}{\cfrac{18}{35}x + \cfrac{1}{\cfrac{7}{6}}}}}}} \quad . \quad . \quad (17)$$

Thus the number of negative α terms is 1 (out of 3), implying 1 pair of conjugate complex zeros and 1 (i.e. $3 - 2 \times 1$) real zero. As the number of negative β terms is also 1 there is no difference in the numbers of negative α and β terms and there are therefore no positive zeros. Hence $f(x)$ has one pair of conjugate complex zeros and one negative zero which may easily be verified, since

$$\begin{aligned} f(x) &= x^3 + 2x^2 + 3x + 2 = (x+1)(x^2 + x + 2) \\ &= (x+1)(x + 1/2 + j\sqrt{7}/2)(x + 1/2 - j\sqrt{7}/2) \quad . \quad (18) \end{aligned}$$

(3.2) Example 2

To illustrate the case of multiple-order zeros, consider

$$\begin{aligned} f(x) &= x^4 + 2x^3 + 5x^2 + 4x + 4 = (x^2 + x + 2)^2 \\ &= [(x + 1/2 + j\sqrt{7}/2)(x + 1/2 - j\sqrt{7}/2)]^2 \quad . \quad (19) \end{aligned}$$

Then
$$f'(x) = 4x^3 + 6x^2 + 10x + 4 \quad . \quad . \quad . \quad (20)$$

and
$$\frac{f(x)}{f'(x)} = \frac{x}{4} + \cfrac{1}{8 + \cfrac{1}{-\cfrac{x}{28} + \cfrac{1}{-7}}} \quad . \quad . \quad . \quad (21)$$

Multiple-order zeros are indicated by premature termination of the continued fraction. Details of these could be obtained,

f required, by inspecting the last divisor in the process of obtaining the continued fraction. By Euclid's algorithm this last divisor is the common factor between $f(x)$ and $f'(x)$.

In the present example, as one negative and one positive x coefficient in the continued fraction indicate one pair of conjugate complex zeros and no real zeros, the conjugate complex zeros must be of order 2.

(3.3) Example 3

Even when the given polynomial is such that a continued-fraction expansion of the same type as in eqn. (3) is no longer obtained, correct results can still be obtained by the following method.*

Suppose, for example,

$$f(x) = x^2 - 1 = (x+1)(x-1) \quad . \quad . \quad . \quad (22)$$

Then

$$\frac{f'(x)}{f(x)} = \frac{2x}{x^2-1} = \frac{1}{\frac{x}{2} + \frac{1}{-2x}} \quad . \quad . \quad . \quad . \quad (23)$$

But, as the continued fraction in eqn. (23) is not of the same type as that in eqn. (3), no conclusions can be drawn with regard to the character of the zeros of $f(x)$; e.g. it would be incorrect to say that one negative α coefficient implied that $f(x)$ contained one pair of conjugate complex zeros.

Now modify $f(x)$ by adding a term ax.

Then

$$\frac{f'(x)}{f(x)} = \frac{2x+a}{x^2+ax-1} = \cfrac{1}{\frac{x}{2} + \cfrac{1}{\frac{4}{a} + \cfrac{1}{\frac{ax}{2\left(a+\frac{4}{a}\right)} + \cfrac{1}{-\left(a+\frac{4}{a}\right)}}}} \quad . \quad . \quad . \quad . \quad (24)$$

Let $a \to 0$ and examine the signs of the coefficients in the continued fraction. Both α terms are positive, indicating no complex zeros and 2 real zeros. Whether $a \to +0$ or -0, there is one negative β term and hence one of the real zeros is positive and the other negative. That this is the correct result is easily verified.

(3.4) Example 4

This example illustrates the determination of the character of the zeros of a polynomial by an examination of sign changes in sequences of determinants.

From the argument connected with sequence (11) of Section 2, if N is the number of changes of sign in the sequence

$$a_n b_{n-1},\ a_n \Delta_3,\ a_n \Delta_5,\ \ldots\ a_n \Delta_{2k-1},\ \ldots\ a_n \Delta_{2n-1} \quad . \quad (25)$$

with Δ_{2k-1} the highest-order non-zero determinant in the sequence, then the given polynomial has k distinct zeros comprising $(k - 2N)$ real zeros and N pairs of conjugate complex zeros.

For convenience in calculation, the Hurwitz type determinants Δ_{2l-1} in sequence (25) can be reduced to about half their original order by using the condensation procedure described for the Hurwitz determinants in a previous paper;[9] some further details are also given in the Appendix.

Thus sequence (25) becomes

$$a_n b_{n-1},\ a_n D_3,\ a_n D_5,\ \ldots\ a_n D_{2k-1},\ \ldots\ a_n D_{2n-1} \quad . \quad (26)$$

* This is one of a number of ways of modifying the polynomial so that the standard form of continued fraction results. Alternatively, the non-standard type of continued fraction originally produced might be investigated.

where $a_n b_{n-1}$ is always positive and D_{2i-1} is the condensed determinant equal to Δ_{2i-1}.

To illustrate this method take the same cubic polynomial used in Example 1, namely $x^3 + 2x^2 + 3x + 2$. Adapting the results for the condensed Hurwitz determinants given in Reference 9, the highest-order determinant to be considered, D_5, is given by

$$D_5 = \begin{vmatrix} a_n b_{n-1} & a_n b_{n-2} & a_n b_{n-3} \\ a_n b_{n-2} & a_n b_{n-3} + \begin{vmatrix} a_{n-1} & a_{n-2} \\ b_{n-1} & b_{n-2} \end{vmatrix} & \begin{vmatrix} a_{n-1} & a_{n-3} \\ b_{n-1} & b_{n-3} \end{vmatrix} \\ a_n b_{n-3} & \begin{vmatrix} a_{n-1} & a_{n-3} \\ b_{n-1} & b_{n-3} \end{vmatrix} & \begin{vmatrix} a_{n-2} & a_{n-3} \\ b_{n-2} & b_{n-3} \end{vmatrix} \end{vmatrix} \quad . \quad (27)$$

D_3 is the second-order principal minor indicated by the dotted lines in the above determinant and $a_n b_{n-1}$ is the first-order principal minor.

For the polynomial $x^3 + 2x^2 + 3x + 2$,

$$\left.\begin{aligned} a_n = 1\ a_{n-1} = 2 \quad a_{n-2} = 3 \quad a_{n-3} = 2 \\ b_{n-1} = 3 \quad b_{n-2} = 4 \quad b_{n-3} = 3 \end{aligned}\right\} \quad . \quad . \quad (28)$$

The terms a_{n-4}, a_{n-5}, etc., and b_{n-4}, b_{n-5}, etc., are zero. Substituting these values in eqn. (27),

$$D_5 = \begin{vmatrix} 3 & 4 & 3 \\ 4 & 3 + \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} & \begin{vmatrix} 2 & 2 \\ 3 & 3 \end{vmatrix} \\ 3 & \begin{vmatrix} 2 & 2 \\ 3 & 3 \end{vmatrix} & \begin{vmatrix} 3 & 2 \\ 4 & 3 \end{vmatrix} \end{vmatrix} = -28 \quad . \quad . \quad . \quad (29)$$

Thus D_5 is negative, indicating one change of sign in sequence (26), and consequently the polynomial $x^3 + 2x^2 + 3x + 2$ has one pair of conjugate complex zeros and one real zero, exactly as before.

The number of positive zeros is given by the excess in the number of sign changes in sequence (9) over those in sequence (26). As before, the determinants in sequence (9) can be condensed to half their original order giving

$$1,\ D_2,\ D_4,\ \ldots\ D_{2k},\ \ldots\ D_{2n} \quad . \quad . \quad . \quad . \quad (30)$$

For the present example the highest-order determinant to be considered, D_6, is given by

$$D_6 = \begin{vmatrix} \begin{vmatrix} b_{n-1} & b_{n-2} \\ a_n & a_{n-1} \end{vmatrix} & \begin{vmatrix} b_{n-1} & b_{n-3} \\ a_n & a_{n-2} \end{vmatrix} & b_{n-1}a_{n-3} \\ \begin{vmatrix} b_{n-1} & b_{n-3} \\ a_n & a_{n-2} \end{vmatrix} & b_{n-1}a_{n-3} + \begin{vmatrix} b_{n-2} & b_{n-3} \\ a_{n-1} & a_{n-2} \end{vmatrix} & b_{n-2}a_{n-3} \\ b_{n-1}a_{n-3} & b_{n-2}a_{n-3} & b_{n-3}a_{n-3} \end{vmatrix} \quad . \quad (31)$$

The second- and first-order principal minors indicated by the dotted lines in the determinant D_6 are equal to D_4 and D_2, respectively.

Substituting the numerical values,

$$D_6 = \begin{vmatrix} 2 & 6 & 6 \\ 6 & 12 & 8 \\ 6 & 8 & 6 \end{vmatrix} = -56 \quad . \quad . \quad . \quad . \quad (32)$$

As D_2 (=2) is positive there is one change of sign in sequence (30), which is the same as the number of changes of sign in sequence (26). The polynomial $x^3 + 2x^2 + 3x + 2$ has, therefore, no positive zeros, agreeing with the conclusion reached in Section 3.1. In this case the result is immediately obvious, since the polynomial, having positive coefficients, can never vanish for positive x.

(4) CONCLUSIONS

It is shown that the number of sign changes along one or two sequences of determinants can be used to determine the number of positive zeros and the numbers of real and complex zeros of a real polynomial. Two theorems are developed to link these sign changes with the number of minus signs occurring in a certain continued fraction. Both results provide useful methods for investigating the character of the zeros of a polynomial. When using the determinantal method it is advantageous to condense the determinants, thus reducing their order to about half. This is illustrated by a numerical example in Section 3.4, and further details are given in the Appendix.

(5) REFERENCES

(1) BURNSIDE, W. S., and PANTON, A. W.: 'Theory of Equations', 3rd edition (Hodges, Figgis and Co., 1892), p. 183.

(2) TURNBULL, H. W.: 'Theory of Equations', 4th edition (Oliver and Boyd, 1947), p. 103.

(3) FULLER, A. T.: 'Conditions for Aperiodicity in Linear Systems', *British Journal of Applied Physics*, 1955, **6**, pp. 195 and 450.

(4) FULLER, A. T.: 'Stability Criteria for Linear Systems and Realizability Criteria for RC Networks', *Proceedings of the Cambridge Philosophical Society*, 1957, **53**, p. 878.

(5) WALL, H. S.: 'Analytic Theory of Continued Fractions' (Van Nostrand, 1948), Chapter 10, p. 173.

(6) TRUDI, N.: 'Teoria de' Determinanti' (Pellerano, Naples 1862), p. 192.

(7) MUIR, T.: 'Theory of Determinants' (Macmillan, 1920) Vol. 3, p. 347.

(8) BADER, W.: 'Beitrag zur Verwirklichung von Wechselstrom widerständen vorgeschriebener Frequenzabhängigkeit' *Archiv für Elektrotechnik*, 1940, **34**, p. 293.

(9) CUTTERIDGE, O. P. D.: 'The Stability Criteria for Linea Systems', *Proceedings I.E.E.*, Monograph No. 328 M February, 1959 (**106** C, p. 125).

(6) APPENDIX

The Appendix discusses the condensation and subsequen simplification of the Hurwitz-type determinants Δ_{2i} and $a_n\Delta_{2i-}$ of sequences (9) and (25) and comments on their significanc in the theory of equations.

Consider first the case of the determinant Δ_{2n} of order $2n$ Adapting the condensation procedure of Reference 9, the follow ing result is finally obtained:

$$
D_{2n} = \begin{array}{|c|c|c|c|c|c|}
\hline
a_n a_{n-1} & 2a_n a_{n-2} & 3a_n a_{n-3} & 4a_n a_{n-4} & 5a_n a_{n-5} & \cdots \\
\hline
2a_n a_{n-2} & \begin{vmatrix} a_{n-1} & -3a_{n-3} \\ a_n & a_{n-2} \end{vmatrix} & \begin{vmatrix} a_{n-1} & -4a_{n-4} \\ a_n & 2a_{n-3} \end{vmatrix} & \begin{vmatrix} a_{n-1} & -5a_{n-5} \\ a_n & 3a_{n-4} \end{vmatrix} & \begin{vmatrix} a_{n-1} & -6a_{n-6} \\ a_n & 4a_{n-5} \end{vmatrix} & \cdots \\
\hline
3a_n a_{n-3} & \begin{vmatrix} a_{n-1} & -4a_{n-4} \\ a_n & 2a_{n-3} \end{vmatrix} & \begin{vmatrix} a_{n-2} & -3a_{n-4} \\ a_{n-1} & a_{n-3} \end{vmatrix} + 5a_n a_{n-5} & \begin{vmatrix} a_{n-2} & -4a_{n-5} \\ a_{n-1} & 2a_{n-4} \end{vmatrix} + 6a_n a_{n-6} & \begin{vmatrix} a_{n-2} & -5a_{n-6} \\ a_{n-1} & 3a_{n-5} \end{vmatrix} + 7a_n a_{n-7} & \cdots \\
\hline
4a_n a_{n-4} & \begin{vmatrix} a_{n-1} & -5a_{n-5} \\ a_n & 3a_{n-4} \end{vmatrix} & \begin{vmatrix} a_{n-2} & -4a_{n-5} \\ a_{n-1} & 2a_{n-4} \end{vmatrix} + 6a_n a_{n-6} & \begin{vmatrix} a_{n-3} & -3a_{n-5} \\ a_{n-2} & a_{n-4} \end{vmatrix} + \begin{vmatrix} a_{n-1} & -7a_{n-7} \\ a_n & 5a_{n-6} \end{vmatrix} & \begin{vmatrix} a_{n-3} & -4a_{n-6} \\ a_{n-2} & 2a_{n-5} \end{vmatrix} + \begin{vmatrix} a_{n-1} & -8a_{n-8} \\ a_n & 6a_{n-7} \end{vmatrix} & \cdots \\
\hline
5a_n a_{n-5} & \begin{vmatrix} a_{n-1} & -6a_{n-6} \\ a_n & 4a_{n-5} \end{vmatrix} & \begin{vmatrix} a_{n-2} & -5a_{n-6} \\ a_{n-1} & 3a_{n-5} \end{vmatrix} + 7a_n a_{n-7} & \begin{vmatrix} a_{n-3} & -4a_{n-6} \\ a_{n-2} & 2a_{n-5} \end{vmatrix} + \begin{vmatrix} a_{n-1} & -8a_{n-8} \\ a_n & 6a_{n-7} \end{vmatrix} & \begin{vmatrix} a_{n-4} & -3a_{n-6} \\ a_{n-3} & a_{n-5} \end{vmatrix} + \begin{vmatrix} a_{n-2} & -7a_{n-8} \\ a_{n-1} & 5a_{n-7} \end{vmatrix} + 9a_n a_{n-9} & \cdots \\
\hline
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\hline
\end{array} \qquad \ldots\ldots (33
$$

where D_{2n} is symmetrical and of order n and equal to Δ_{2n}. Values of D_{2n} for $n = 3$, 4, 5 and 6 are

$$
n = 3: \quad \begin{array}{|c|c|c|}
\hline
a_3 a_2 & 2a_3 a_1 & 3a_3 a_0 \\
\hline
2a_3 a_1 & \begin{vmatrix} a_2 & -3a_0 \\ a_3 & a_1 \end{vmatrix} & 2a_2 a_0 \\
\hline
3a_3 a_0 & 2a_2 a_0 & a_1 a_0 \\
\hline
\end{array}
$$

$$
n = 4: \quad \begin{array}{|c|c|c|c|}
\hline
a_4 a_3 & 2a_4 a_2 & 3a_4 a_1 & 4a_4 a_0 \\
\hline
2a_4 a_2 & \begin{vmatrix} a_3 & -3a_1 \\ a_4 & a_2 \end{vmatrix} & \begin{vmatrix} a_3 & -4a_0 \\ a_4 & 2a_1 \end{vmatrix} & 3a_3 a_0 \\
\hline
3a_4 a_1 & \begin{vmatrix} a_3 & -4a_0 \\ a_4 & 2a_1 \end{vmatrix} & \begin{vmatrix} a_2 & -3a_0 \\ a_3 & a_1 \end{vmatrix} & 2a_2 a_0 \\
\hline
4a_4 a_0 & 3a_3 a_0 & 2a_2 a_0 & a_1 a_0 \\
\hline
\end{array}
$$

$$
n = 5: \quad \begin{array}{|c|c|c|c|c|}
\hline
a_5 a_4 & 2a_5 a_3 & 3a_5 a_2 & 4a_5 a_1 & 5a_5 a_0 \\
\hline
2a_5 a_3 & \begin{vmatrix} a_4 & -3a_2 \\ a_5 & a_3 \end{vmatrix} & \begin{vmatrix} a_4 & -4a_1 \\ a_5 & 2a_2 \end{vmatrix} & \begin{vmatrix} a_4 & -5a_0 \\ a_5 & 3a_1 \end{vmatrix} & 4a_4 a_0 \\
\hline
3a_5 a_2 & \begin{vmatrix} a_4 & -4a_1 \\ a_5 & 2a_2 \end{vmatrix} & \begin{vmatrix} a_3 & -3a_1 \\ a_4 & a_2 \end{vmatrix} + 5a_5 a_0 & \begin{vmatrix} a_3 & -4a_0 \\ a_4 & 2a_1 \end{vmatrix} & 3a_3 a_0 \\
\hline
4a_5 a_1 & \begin{vmatrix} a_4 & -5a_0 \\ a_5 & 3a_1 \end{vmatrix} & \begin{vmatrix} a_3 & -4a_0 \\ a_4 & 2a_1 \end{vmatrix} & \begin{vmatrix} a_2 & -3a_0 \\ a_3 & a_1 \end{vmatrix} & 2a_2 a_0 \\
\hline
5a_5 a_0 & 4a_4 a_0 & 3a_3 a_0 & 2a_2 a_0 & a_1 a_0 \\
\hline
\end{array}
$$

$$
n = 6: \quad \begin{array}{|c|c|c|c|c|c|}
\hline
a_6 a_5 & 2a_6 a_4 & 3a_6 a_3 & 4a_6 a_2 & 5a_6 a_1 & 6a_6 a \\
\hline
2a_6 a_4 & \begin{vmatrix} a_5 & -3a_3 \\ a_6 & a_4 \end{vmatrix} & \begin{vmatrix} a_5 & -4a_2 \\ a_6 & 2a_3 \end{vmatrix} & \begin{vmatrix} a_5 & -5a_1 \\ a_6 & 3a_2 \end{vmatrix} & \begin{vmatrix} a_5 & -6a_0 \\ a_6 & 4a_1 \end{vmatrix} & 5a_5 a \\
\hline
3a_6 a_3 & \begin{vmatrix} a_5 & -4a_2 \\ a_6 & 2a_3 \end{vmatrix} & \begin{vmatrix} a_4 & -3a_2 \\ a_5 & a_3 \end{vmatrix} + 5a_6 a_1 & \begin{vmatrix} a_4 & -4a_1 \\ a_5 & 2a_2 \end{vmatrix} + 6a_6 a_0 & \begin{vmatrix} a_4 & -5a_0 \\ a_5 & 3a_1 \end{vmatrix} & 4a_4 a \\
\hline
4a_6 a_2 & \begin{vmatrix} a_5 & -5a_1 \\ a_6 & 3a_2 \end{vmatrix} & \begin{vmatrix} a_4 & -4a_1 \\ a_5 & 2a_2 \end{vmatrix} + 6a_6 a_0 & \begin{vmatrix} a_3 & -3a_1 \\ a_4 & a_2 \end{vmatrix} + 5a_5 a_0 & \begin{vmatrix} a_3 & -4a_0 \\ a_4 & 2a_1 \end{vmatrix} & 3a_3 a \\
\hline
5a_6 a_1 & \begin{vmatrix} a_5 & -6a_0 \\ a_6 & 4a_1 \end{vmatrix} & \begin{vmatrix} a_4 & -5a_0 \\ a_5 & 3a_1 \end{vmatrix} & \begin{vmatrix} a_3 & -4a_0 \\ a_4 & 2a_1 \end{vmatrix} & \begin{vmatrix} a_2 & -3a_0 \\ a_3 & a_1 \end{vmatrix} & 2a_2 a \\
\hline
6a_6 a_0 & 5a_5 a_0 & 4a_4 a_0 & 3a_3 a_0 & 2a_2 a_0 & a_1 a \\
\hline
\end{array}
$$

For each of the values of n the determinant given is the highest-order determinant in sequence (30); the minors indicated by the dotted lines are the remaining determinantal members of sequence (30) in each case.

Consider now the determinant $a_n\Delta_{2n-1}$. This is of order $(2n-1)$ and is reduced to the following symmetrical determinant, a_nD_{2n-1}, of order n, by applying the appropriate condensation procedure from Reference 9 and simplifying the result:

As before, each determinant given above is the highest-order determinant in sequence (26) and the minors indicated by the dotted lines are in each case the remaining determinantal members of sequence (26).

Now the determinants D_{2n} and a_nD_{2n-1} can be shown to be simply related.

$$a_nD_{2n-1} =$$

$$\begin{array}{llllll}
na_n^2 & (n-1)a_na_{n-1} & (n-2)a_na_{n-2} & (n-3)a_na_{n-3} & (n-4)a_na_{n-4} & \cdots \\
-1)a_na_{n-1} & \begin{vmatrix} a_{n-1} & 2a_{n-2} \\ a_n & (n-1)a_{n-1} \end{vmatrix} & \begin{vmatrix} a_{n-1} & 3a_{n-3} \\ a_n & (n-2)a_{n-2} \end{vmatrix} & \begin{vmatrix} a_{n-1} & 4a_{n-4} \\ a_n & (n-3)a_{n-3} \end{vmatrix} & \begin{vmatrix} a_{n-1} & 5a_{n-5} \\ a_n & (n-4)a_{n-4} \end{vmatrix} & \cdots \\
-2)a_na_{n-2} & \begin{vmatrix} a_{n-1} & 3a_{n-3} \\ a_n & (n-2)a_{n-2} \end{vmatrix} & \begin{vmatrix} a_{n-2} & 2a_{n-3} \\ a_{n-1} & (n-2)a_{n-2} \end{vmatrix} - 4a_na_{n-4} & \begin{vmatrix} a_{n-2} & 3a_{n-4} \\ a_{n-1} & (n-3)a_{n-3} \end{vmatrix} - 5a_na_{n-5} & \begin{vmatrix} a_{n-2} & 4a_{n-5} \\ a_{n-1} & (n-4)a_{n-4} \end{vmatrix} - 6a_na_{n-6} & \cdots \\
-3)a_na_{n-3} & \begin{vmatrix} a_{n-1} & 4a_{n-4} \\ a_n & (n-3)a_{n-3} \end{vmatrix} & \begin{vmatrix} a_n\ 2 & 3a_{n-4} \\ a_{n-1} & (n-3)a_{n-3} \end{vmatrix} - 5a_na_{n-5} & \begin{vmatrix} a_{n-3} & 2a_{n-4} \\ a_{n-2} & (n-3)a_{n-3} \end{vmatrix} + \begin{vmatrix} a_{n-1} & 6a_{n-6} \\ a_n & -4a_{n-5} \end{vmatrix} & \begin{vmatrix} a_{n-3} & 3a_{n-5} \\ a_{n-2} & (n-4)a_{n-4} \end{vmatrix} + \begin{vmatrix} a_n\ {\cdot 1} & 7a_{n-7} \\ a_n & -5a_{n-6} \end{vmatrix} & \cdots \\
-4)a_na_{n-4} & \begin{vmatrix} a_{n-1} & 5a_{n-5} \\ a_n & (n-4)a_{n-4} \end{vmatrix} & \begin{vmatrix} a_{n-2} & 4a_{n-5} \\ a_{n-1} & (n-4)a_{n-4} \end{vmatrix} - 6a_na_{n-6} & \begin{vmatrix} a_{n-3} & 3a_{n-5} \\ a_{n-2} & (n-4)a_{n-4} \end{vmatrix} + \begin{vmatrix} a_{n-1} & 7a_{n-7} \\ a_n & -5a_{n-6} \end{vmatrix} & \begin{vmatrix} a_{n-4} & 2a_{n-5} \\ a_{n-3} & (n-4)a_{n-4} \end{vmatrix} + \begin{vmatrix} a_n\ {-1} & 8a_{n-8} \\ a_n & -6a_{n-7} \end{vmatrix} - 4a_{n-2}a_{n-6} & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots
\end{array} \quad (34)$$

Values of a_nD_{2n-1} for $n = 3$, 4 and 5 are

$n = 3$:

$$\begin{vmatrix}
3a_3^2 & 2a_3a_2 & a_3a_1 \\
2a_3a_2 & \begin{vmatrix} a_2 & 2a_1 \\ a_3 & 2a_2 \end{vmatrix} & \begin{vmatrix} a_2 & 3a_0 \\ a_3 & a_1 \end{vmatrix} \\
a_3a_1 & \begin{vmatrix} a_2 & 3a_0 \\ a_3 & a_1 \end{vmatrix} & \begin{vmatrix} a_1 & 2a_0 \\ a_2 & a_1 \end{vmatrix}
\end{vmatrix}$$

$n = 4$:

$$\begin{vmatrix}
4a_4^2 & 3a_4a_3 & 2a_4a_2 & a_4a_1 \\
3a_4a_3 & \begin{vmatrix} a_3 & 2a_2 \\ a_4 & 3a_3 \end{vmatrix} & \begin{vmatrix} a_3 & 3a_1 \\ a_4 & 2a_2 \end{vmatrix} & \begin{vmatrix} a_3 & 4a_0 \\ a_4 & a_1 \end{vmatrix} \\
2a_4a_2 & \begin{vmatrix} a_3 & 3a_1 \\ a_4 & 2a_2 \end{vmatrix} & \begin{vmatrix} a_2 & 2a_1 \\ a_3 & 2a_2 \end{vmatrix} - 4a_4a_0 & \begin{vmatrix} a_2 & 3a_0 \\ a_3 & a_1 \end{vmatrix} \\
a_4a_1 & \begin{vmatrix} a_3 & 4a_0 \\ a_4 & a_1 \end{vmatrix} & \begin{vmatrix} a_2 & 3a_0 \\ a_3 & a_1 \end{vmatrix} & \begin{vmatrix} a_1 & 2a_0 \\ a_2 & a_1 \end{vmatrix}
\end{vmatrix}$$

$n = 5$:

$$\begin{vmatrix}
5a_5^2 & 4a_5a_4 & 3a_5a_3 & 2a_5a_2 & a_5a_1 \\
4a_5a_4 & \begin{vmatrix} a_4 & 2a_3 \\ a_5 & 4a_4 \end{vmatrix} & \begin{vmatrix} a_4 & 3a_2 \\ a_5 & 3a_3 \end{vmatrix} & \begin{vmatrix} a_4 & 4a_1 \\ a_5 & 2a_2 \end{vmatrix} & \begin{vmatrix} a_4 & 5a_0 \\ a_5 & a_1 \end{vmatrix} \\
3a_5a_3 & \begin{vmatrix} a_4 & 3a_2 \\ a_5 & 3a_3 \end{vmatrix} & \begin{vmatrix} a_3 & 2a_2 \\ a_4 & 3a_3 \end{vmatrix} - 4a_5a_1 & \begin{vmatrix} a_3 & 3a_1 \\ a_4 & 2a_2 \end{vmatrix} - 5a_5a_0 & \begin{vmatrix} a_3 & 4a_0 \\ a_4 & a_1 \end{vmatrix} \\
2a_5a_2 & \begin{vmatrix} a_4 & 4a_1 \\ a_5 & 2a_2 \end{vmatrix} & \begin{vmatrix} a_3 & 3a_1 \\ a_4 & 2a_2 \end{vmatrix} - 5a_5a_0 & \begin{vmatrix} a_2 & 2a_1 \\ a_3 & 2a_2 \end{vmatrix} - 4a_4a_0 & \begin{vmatrix} a_2 & 3a_0 \\ a_3 & a_1 \end{vmatrix} \\
a_5a_1 & \begin{vmatrix} a_4 & 5a_0 \\ a_5 & a_1 \end{vmatrix} & \begin{vmatrix} a_3 & 4a_0 \\ a_4 & a_1 \end{vmatrix} & \begin{vmatrix} a_2 & 3a_0 \\ a_3 & a_1 \end{vmatrix} & \begin{vmatrix} a_1 & 2a_0 \\ a_2 & a_1 \end{vmatrix}
\end{vmatrix}$$

For $n = 3$,

$$a_3D_5 = \left(\frac{a_3}{a_0}\right)D_6 \quad . \quad . \quad . \quad . \quad . \quad (35)$$

For $n = 4$,

$$a_4D_7 = \left(\frac{a_4}{a_0}\right)D_8 \quad . \quad . \quad . \quad . \quad . \quad (36)$$

and, in general,

$$a_nD_{2n-1} = \left(\frac{a_n}{a_0}\right)D_{2n} \quad . \quad . \quad . \quad . \quad . \quad (37)$$

A proof of this relationship for $n = 3$ is as follows. In the determinant a_nD_{2n-1} for $n = 3$, take

2nd column $-$ $(a_2/a_3) \times$ 1st column,
3rd column $-$ $(a_1/a_3) \times$ 1st column

and take out a factor a_3/a_0 from the 1st column. On interchanging columns 1 and 2 and then columns 2 and 3, and multiplying columns 1 and 2 by -1 we have the value of $(a_n/a_0)D_{2n}$ for $n = 3$.

This simple relationship between the two highest-order determinants in each sequence means that it is sufficient to consider only one such determinant, and all the determinants in both sequences are either equal or simply related to the minors of this one determinant. Since the determinants D_{2n} are seen to be rather simpler than the determinants a_nD_{2n-1} it is convenient to take the D_{2n} as basic and to refer the others to this. Thus, for example, if D is the value of D_{2n} when $n = 3$, sequence (30) can be written

$$1, D_{3322}, D_{33}, D \quad . \quad . \quad . \quad . \quad . \quad (38)$$

and sequence (26) can be written

$$\frac{a_3}{a_0}D_{3221}, \frac{a_3}{a_0}D_{32}, \frac{a_3}{a_0}D \quad . \quad . \quad . \quad . \quad . \quad (39)$$

where D_{32}, D_{3221} are minors,* not cofactors.

* The notation used here is that the first and third suffixes indicate rows and the second and fourth indicate columns. With this notation the result corresponding to sequence (39) for the case $n = 4$ is a_4D_{433221}/a_0, a_4D_{4332}/a_0, a_4D_{43}/a_0, a_4D/a_0, where D is now the value of D_{2n} when $n = 4$. Similarly for higher-order cases.

Writing Δ for the last term $[=(a_n/a_0)D_{2n}]$ in sequence (26), it is seen that the sign of Δ has a great significance on the possible character of the zeros of the polynomial. Since the first term of sequence (26) $(= na_n^2)$ must always be positive, $\Delta > 0$ indicates 0, 2, 4, 6, . . . sign changes in sequence (26), i.e.

$\Delta > 0$ indicates 0, 4, 8, 12, . . . complex zeros in the polynomial . . . (40)

and $\Delta < 0$ indicates 1, 3, 5, 7, . . . sign changes in sequence (26), i.e.

$\Delta < 0$ indicates 2, 6, 10, 14, . . . complex zeros in the polynomial . . . (41)

Δ might therefore be described as the discriminant of the polynomial, and it is in fact proportional to the usual form of the discriminant in the cases of the quadratic, cubic and quartic. In the case of the cubic, for example, Δ is equal to -108 times the usual discriminant. The above results therefore provide a generalization of the idea of the discriminant of a polynomial for the quintic and higher orders.

The following observations can be drawn from statements (40) and (41). For the quadratic and cubic the sign of the discriminant is always sufficient to determine exactly how many zeros are complex; in each case $\Delta < 0$ indicates two complex zeros and $\Delta > 0$ indicates that all zeros are real. For the quartic and quintic, if $\Delta < 0$ it is seen that there are exactly two complex zeros, whereas if $\Delta > 0$ there is ambiguity in that there may be either 0 or 4 complex zeros in each case. In the latter case the signs of the appropriate minors must also be determined in order to overcome this ambiguity. For polynomials of the 6th degree and higher the sign of the discriminant alone is never sufficient to determine exactly how many zeros are real and how many complex. The above observations on the quadratic, cubic and quartic correspond to well-known results (see, for example, References 1 and 2) in the theory of equations.

DISCUSSION ON 'THE STABILITY CRITERIA FOR LINEAR SYSTEMS'*

Mr. A. T. Fuller (*communicated*): It would be interesting to know if Dr. Cutteridge was led to his condensed determinants from a consideration of Bezoutian determinants, which occur in the theory of elimination.[A]

I have recently come across some German work on the stability criteria. In 1953 L. Cremer[B] showed that these may be reduced to the requirement that (if $a_n > 0$)

$$\Delta_{n-1}, \Delta_{n-3}, \Delta_{n-5} \quad . \quad . \quad . \quad . \quad . \quad . \quad (A)$$

and

$$a_0, a_1, a_2, \ldots, a_{n-1} \quad . \quad . \quad . \quad . \quad . \quad . \quad (B)$$

should all be positive. Comparison with the results given by Dr. Cutteridge and by myself[C] shows that some of the coefficients in sequence (B) are redundant. Incidentally, L. Cremer also proved that sequence (A) may be replaced by

$$\Delta_n, \Delta_{n-2}, \Delta_{n-4}, \ldots \quad . \quad . \quad . \quad . \quad . \quad (C)$$

In 1957 Effertz[D] obtained condensed determinants similar to those of Dr. Cutteridge. However, Effertz's proof depends on a paper[E] which has apparently not yet been published. Effertz's criteria include redundant members, in that they require all sequence (B) to be positive. Effertz has also treated[F,G] the network proof of Bückner's canonical form.[H,I]

Mr. J. W. Head (*communicated*): While Dr. Cutteridge is to be congratulated on the neat way in which he has welded together various alternative stability criteria and shown the connection between them, I would like to suggest that far the most important of these criteria is that the zeros of the even and odd parts of a Hurwitz polynomial are simple, purely imaginary and separate one another, and that this criterion is much more directly useful than is generally supposed.

There are many practical cases in which we have to deal with a polynomial $f(p)$ of degree not exceeding 7, so that the polynomial

$$f_7(p) = p^7 + a_6p^6 + a_5p^5 + a_4p^4 + a_3p^3 + a_2p^2 + a_1p + a_0$$
$$= p(p^2 + \beta_1)(p^2 + \beta_2)(p^2 + \beta_3) + a_6(p^2 + x_1)(p^2 + x_2)(p^2 + x_3)$$

will provide a good illustration. If $f_7(p)$ is a Hurwitz polynomial, all the α's and β's will be real and positive and will separate each other. If we are given the a's and have to determine the α's and β's and test for stability, we have only to determine whether two cubics in p^2 have all their roots real. Only if the two cubics both have all their roots real do we have to consider the question whether they separate each other.

It is relatively easy to find the real roots of an algebraic equation (especially when they are not close together) either graphically or by some successive approximation process such as that due to Lin. If the equation under consideration is a cubic it must have at least one real root, and when this is found and the corresponding factor is divided out, only a quadratic remains, so that it becomes immediately obvious whether the remaining pair of roots is real or complex.

If the coefficients of the characteristic polynomial [such as $f_7(p)$ above] contain a variable parameter, the zeros of the even and odd parts can be obtained for a sufficient number of values of that parameter, and if these zeros are plotted against the variable parameter, the way in which stability breaks down at critical values of the parameter becomes clear.

Dr. O. P. D. Cutteridge (*in reply*): I believe I was led to the condensed determinants after studying various condensation techniques given in a number of textbooks. The relationship of the condensed determinants to the Bezoutians is the same as that of the Hurwitz determinants to the bigradients[J] of Sylvester's dialytic method of elimination.

* Cutteridge, O. P. D.: Monograph No. 328 M February, 1959 (see 106 C, p. 125).

I would agree with Mr. Fuller's comments on the German work that he quotes. Incidentally, an account in English of some of this and related material is to be found in a section contributed by Effertz to the recent translation of Cauer's treatise.[K] The paper by Cremer and Effertz[L] was published in March, 1959, and contains a total of 38 references; the method of treatment differs completely from that adopted in my paper.

It is interesting to note that a canonical form for a Hurwitz polynomial can be obtained[M] as a determinant whose order is about half that of Bückner's form. For example, for the sixth-degree case, the expression

$$p^3 \begin{vmatrix} 1 + a_1 p + \dfrac{b_1}{p} & -1 & . \\ 1 & b_2 p + \dfrac{1}{a_2 p} & -1 \\ . & 1 & a_3 p + \dfrac{b_3}{p} \end{vmatrix}$$

will be a Hurwitz polynomial for all positive values of the coefficients a_1, a_2, a_3, b_1, b_2, b_3; and conversely any sixth-degree Hurwitz polynomial can be expressed in this form. The third-order determinant above compares with the sixth-order determinant for the corresponding case in Bückner's form.

The question of alternative formulations of the sets of stability criteria, without the introduction of redundant conditions, is of some practical importance and was briefly considered at the end of Appendix 7.3 of the paper. Some further variations will be given in a forthcoming paper.[M]

Incidentally the basic result that either the Hurwitz determinants of even order or those of odd order can be replaced by either of two non-redundant sets of the coefficients was given in a much neglected paper by Liénard and Chipart[N] in 1914 by means of an analysis based on certain quadratic forms associated with the polynomial. These authors were apparently unaware of Hurwitz's paper[O] but compared their results with those of Routh,[P,Q] and with related results of Hermite. The last chapter of a Russian textbook 'Theory of Matrices' by Gantmacher[R] is devoted to stability and related problems, and contains a much more explicit account[S] of the results of Liénard and Chipart than was given by the original authors, together with an alternative proof and some extensions to the basic result. At the moment of writing apparently only the first part of this textbook is available in German translation,[T] but there is an English translation of the relevant second part.[U] The paper by Liénard and Chipart, which unfortunately has only just come to my notice, also gave an example of a condensed determinant, similar to those of my paper, obtained by forming the eliminant of two polynomials by Bezout's method.

I am not sure I agree with Mr. Head in maintaining the pre-eminence of any one form of the stability criterion, from the standpoint of practical application, in view of the varied nature of the problems that can arise from practice.

From the theoretical point of view, however, as the fact that the zeros of the odd and even parts of a Hurwitz polynomial must be simple, purely imaginary and must separate one another is easily obtained, this makes it a very suitable starting-point from which to derive other forms of stability criteria. This was the method adopted by the author and also by Routh,[V] as mentioned in the paper. Routh evidently regarded this form as unsuitable for most applications and proceeded to develop his two well-known rules.

As a general principle it seems preferable to investigate the character of the zeros of a polynomial indirectly rather than calculate them all approximately and then examine the results obtained. Although the polynomials whose zeros have to be determined in Mr. Head's method are only about half the order of the original polynomial, I doubt whether, generally speaking, his method would result in the expenditure of less labour than others.

References

(A) Muir, T.: 'History of determinants' (London: Macmillan, 1920), Vol. 3, p. 328.

(B) Cremer, L.: 'Die Verringerung der Zahl der Stabilitätskriterien bei Voraussetzung positiver Koeffizienten der charakteristischen Gleichung', *Zeitschrift für angewandte Mathematik und Mechanik*, 1953, **33**, p. 221.

(C) Reference 14 of the paper.

(D) Effertz, F. H.: 'Ein Verfahren zur Berechung der Stabilitätsbedingungen bei Voraussetzung positiver Koeffizienten der charakteristischen Gleichung', *Zeitschrift für angewandte Mathematik und Mechanik*, 1957, **37**, p. 487.

(E) Cremer, H., and Effertz, F. H.: 'Über die algebraischen Kriterien für die Stabilität von Schwingungssystemen', *Mathematische Annalen* (in the press).

(F) Effertz, F. H.: 'Beschränkte Funktionen, Frequenzcharakteristiken elektrischer Netzwerke und algebraische Stabilitätskriterien', *Zeitschrift für angewandte Mathematik und Mechanik*, 1953, **33**, p. 281.

(G) Effertz, F. H.: 'On the synthesis of networks containing two kinds of elements', *Proceedings of the Symposium on Modern Network Synthesis*, New York, 1955, p. 145.

(H) Reference 11 of the paper.

(I) Reference 12 of the paper.

(J) Aitken, A. C.: 'Determinants and Matrices', 3rd edn. (Oliver and Boyd, 1944), Section 52, p. 125.

(K) Cauer, W.: 'Synthesis of Linear Communication Networks' (McGraw-Hill, 1958), Vols. 1 and 2, Appendix 5, Section 7, p. 840.

(L) Cremer, H., and Effertz, F. H.: 'Über die algebraischen Kriterien für die Stabilität von Regelungssystemen', *Mathematische Annalen*, 1959, **137**, p. 328.

(M) Cutteridge, O. P. D.: 'Two-terminal RC Networks and Theoretically Related Topics' (in the press).

(N) Liénard and Chipart: 'Sur le signe de la partie réelle des racines d'une équation algébrique', *Journal de Mathématiques pures et appliquées*, 1914, **10** (6e série), p. 291.

(O) Reference 3 of the paper.

(P) Reference 1 of the paper.

(Q) Reference 2 of the paper.

(R) Gantmacher, F. R.: 'Theory of Matrices' (Moscow, 1954), pp. 419–480 (in Russian).

(S) *Ibid.*, p. 457.

(T) Gantmacher, F. R.: 'Matrizenrechnung I', (V.E.B. Deutscher Verlag der Wissenschaften, Berlin, 1958).

(U) Gantmacher, F. R.: 'Applications of the Theory of Matrices' (Interscience Publishers, New York, 1959).

(V) Routh, E. J.: 'A Treatise on the Stability of a Given State of Motion' (Macmillan, 1877), p. 25.

621.319.42

The Institution of Electrical Engineer
Monograph No. 351 M
Jan. 196
©

SOME RESULTS ON THE CROSS-CAPACITANCES PER UNIT LENGTH OF CYLINDRICAL THREE-TERMINAL CAPACITORS WITH THIN DIELECTRIC FILMS ON THEIR ELECTRODES

By D. G. LAMPARD, M.Sc., Ph.D., and R. D. CUTKOSKY, B.Sc.

(*The paper was first received* 1*st April, in revised form* 17*th July, and in final form* 18*th August*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in January*, 1960.)

SUMMARY

The effect on the cross-capacitances per unit length of cylindrical 3-terminal capacitors of thin dielectric films on the capacitor electrodes is discussed.

It is assumed that the cross-section of such dielectric films remains constant throughout the length of the capacitor. Some conformal transformations of basic cylindrical capacitor cross-sections are given. When these are applied to a symmetric cylindrical capacitor with a thin dielectric film on its electrodes, the results suggest that the mean capacitance per unit length may remain constant to the first order despite the presence of the dielectric film. The same methods also suggest that the individual cross-capacitances per unit length may remain constant to the first order provided that the dielectric film is disposed symmetrically with respect to the capacitor symmetry plane.

Further support for these conjectures is given by the results of a detailed calculation of the cross-capacitances per unit length of a parallel-plate cylindrical capacitor with a thin uniform dielectric film on one electrode. In the last Section, a lemma concerning the existence of an equivalent dielectricless capacitor is given and this is followed by the proof of general results of the type suggested by the previous working.

(1) INTRODUCTION

A new approach to the design of calculable 3-terminal capacitance standards has recently been made possible as a consequence of a general theorem[1] concerning cross-capacitances of certain cylindrical electrode systems. This theorem enables a capacitor to be constructed such that only one precision length measurement is necessary to compute its capacitance to a high order of accuracy.

A proof of this general theorem, together with detailed calculations of cross-capacitances of cylindrical systems with a variety of cross-sections, has been given in a recent paper.[2] In that paper the effects of various perturbations which could occur in practice were considered, but an important practical perturbation, namely the presence of spurious thin dielectric films on the capacitor electrodes, was dealt with only in one special case.

Such dielectric films can be present, for example, in a precision calculable cylindrical capacitor, where it may be difficult to avoid surface contamination of the electrodes with oil films, adsorbed vapours, etc. They can also arise in connection with the direct experimental determination (by capacitance measurement of a cylindrical capacitor) of the permittivity of vapours, where a thin film of condensed vapour on the capacitor electrodes may constitute a serious source of error.

In the paper we examine further the effects of thin dielectric films in cylindrical capacitors, the assumption being made that the cross-section of the dielectric film remains constant throughout the length of the capacitor. To facilitate the discussion of these dielectric effects, a number of conformal transformations are given in Section 2 by means of which a cylindrical capacitor of arbitrary cross-section can be mapped into a cylindrical parallel-plate capacitor. This procedure provides a simple alternative proof of the main theorem on cylindrical capacitors in a more generalized form.

In Section 3 it is shown how some uniform dielectric distributions behave under the transformations of Section 2, while in Section 4 both cross-capacitances per unit length of the cylindrical parallel-plate capacitor of Section 2, with a highly asymmetric dielectric distribution on its electrodes, are calculated in detail. The calculations of Section 3 and 4 suggest some general results, which are established in Section 5, concerning the effect of thin dielectric films.

(2) GENERALIZATION OF THE THEOREM ON CYLINDRICAL CAPACITORS

Consider a conducting cylindrical shell whose right cross-section, Fig. 1(*a*), is the arbitrary closed curve S. Let this shell be divided into four parts by infinitesimal gaps at α, β, γ and δ.

By means of an existence theorem due to Riemann,[3] in the theory of conformal transformations, the interior of the cross-section S can be mapped into the interior of a circle, Fig. 1(*b*), and with a suitable choice of a co-ordinate system the point α can be mapped on to the origin and the centre of the circle taken on the real axis.

The mapping $\omega = 1/z$ transforms the interior of the circle of Fig. 1(*b*) onto the right-hand half-plane of Fig. 1(*c*), where the distances $\beta\gamma$ and $\gamma\delta$ have been arbitrarily called b and a, respectively.

By means of simple rotations and translations, the right half-plane of Fig. 1(*c*) can be mapped into the upper half-plane of either Fig. 1(*d*)(i) or (ii).

Finally, the mapping $v = \log_e u$ transforms the upper half-planes of Fig. 1(*d*) into the strip spaces of Fig. 1(*e*).

The configuration of Fig. 1(*e*)(i) is that of the usual guarded parallel-plate capacitor. The capacitance per unit length can be computed by assuming electrode $\alpha\beta$ to be at unit potential and all other electrodes at zero potential. In this case, all lines of force are parallel to the imaginary axis and the capacitance per unit length, C_1, between electrodes $\alpha\beta$ and $\gamma\delta$ is simply

$$C_1 = \frac{1}{4\pi} \frac{\log_\varepsilon (a+b) - \log_\varepsilon b}{\pi} \quad . \quad . \quad . \quad (1)$$

$$= \frac{1}{4\pi^2} \log_\varepsilon \left(1 + \frac{a}{b}\right) \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Similarly, the capacitance per unit length, C_2, between electrodes $\alpha\delta$ and $\beta\gamma$ is,* from Fig. 1(*e*)(ii),

$$C_2 = \frac{1}{4\pi^2} \log_\varepsilon \left(1 + \frac{b}{a}\right) \quad . \quad . \quad . \quad . \quad (3)$$

* Note that eqns. (2) and (3) agree with eqn. (75) of Reference 2.

Correspondence on Monographs is invited for consideration with a view to publication.

Dr. Lampard is in the Division of Electrotechnology of the Commonwealth Scientific and Industrial Research Organization, Australia, and Mr. Cutkosky is at the National Bureau of Standards, Washington, D.C., United States.

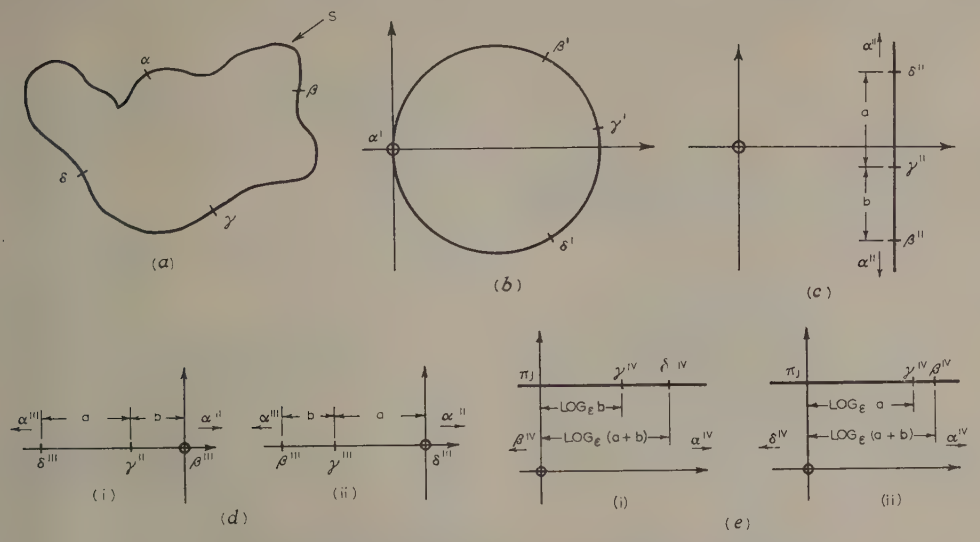


Fig. 1.—Transformed cross-sections of general cylindrical capacitor.

(*a*) Cross-section.
(*b*) Transformation into *z*-plane.
(*c*) Transformation into *w*-plane.
(*d*) Transformation into *u*-plane.
(i) Orientation 1. (ii) Orientation 2.
(*e*) Transformation into *v*-plane.
(i) Orientation 1. (ii) Orientation 2.

If we introduce a parameter $\tau = a/b$ we see that the two cross-capacitances per unit length for any cylindrical capacitor can always be written in the form

$$C_1(\tau) = \frac{1}{4\pi^2}\log_\varepsilon(1+\tau) \quad . \quad . \quad . \quad (4)$$

$$C_2(\tau) = \frac{1}{4\pi^2}\log_\varepsilon\left(1+\frac{1}{\tau}\right) \quad . \quad . \quad . \quad . \quad (5)$$

Eliminating τ between these two equations,

$$\varepsilon^{-4\pi^2 C_1} + \varepsilon^{-4\pi^2 C_2} = 1 \quad . \quad . \quad . \quad . \quad (6)$$

This identity was given as eqns. (32) and (71) of Reference 2 in connection with cylindrical capacitors of square and circular cross-section, respectively, but its complete generality was not noticed until some time later (Lampard[4] and more recently van der Pauw[5]).

It follows immediately from eqn. (6) that, if $C_1 = C_2$,

$$C_1 = C_2 = \frac{1}{4\pi^2}\log_\varepsilon 2 \quad . \quad . \quad . \quad . \quad . \quad (7)$$

It is clear that a sufficient, but not necessary, condition for the equality of C_1 and C_2 is the existence of an axis of symmetry of the capacitor cross-section, which passes through a pair of non-adjacent electrode gaps. This was the form in which the theorem was originally stated. It will be clear that the cross-section of any cylindrical capacitor whose cross-capacitances per unit length are equal can always be mapped into a cross-section having this symmetry property, and hence, in the paper, any such cylindrical capacitor will be referred to as *electrically symmetric*. The non-necessity of the physical symmetry condition has been pointed out by Thompson.[6]

On differentiating eqn. (6),

$$\varepsilon^{-4\pi^2 C_1} dC_1 = -\varepsilon^{-4\pi^2 C_2} dC_2 \quad . \quad . \quad . \quad (8)$$

and, in particular, when $C_1 = C_2$,

$$dC_1 = -dC_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

We shall make use of this result in a later Section to prove some general results concerning the effects of thin dielectric films.

(3) CONFORMAL MAPPING OF DIELECTRIC DISTRIBUTIONS

The way in which two simple uniform dielectric distributions behave under the conformal transformations of Section 2 will be examined. Thus consider the physically symmetrical system of Fig. 2(*a*). This system was analysed in Reference 2, where

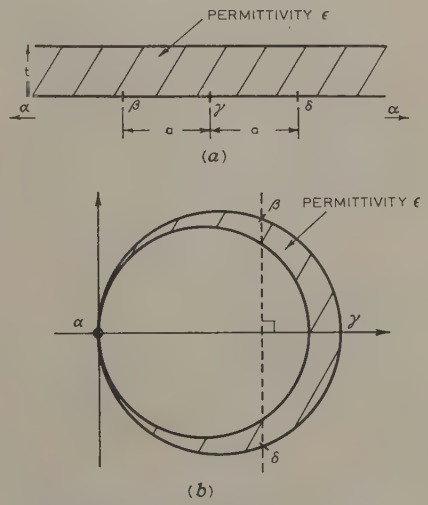


Fig. 2.—Infinite-plane cylindrical capacitor with dielectric film.

(*a*) Cross-section.
(*b*) Transformation into *z*-plane.

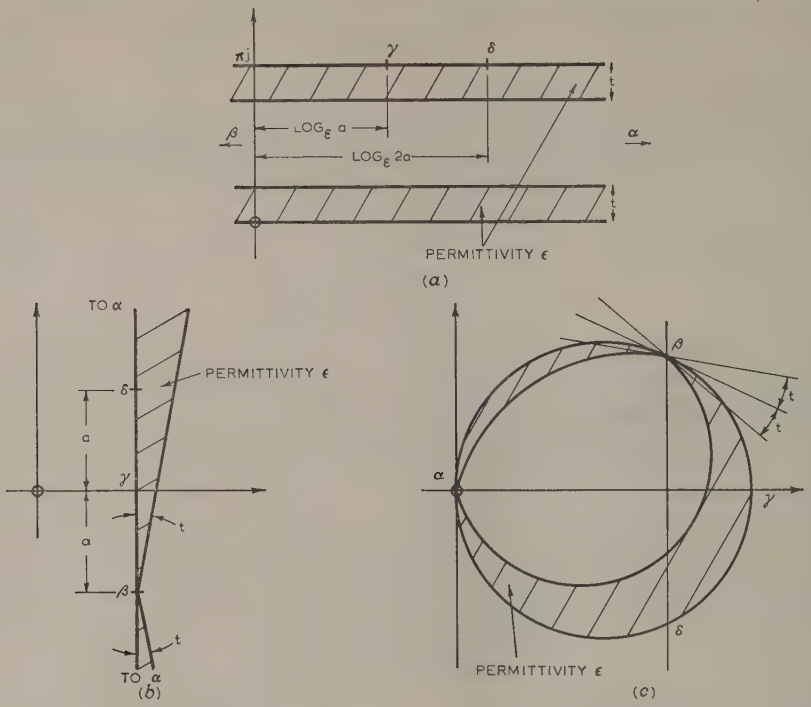


Fig. 3.—Parallel-plate cylindrical capacitor with dielectric film.

(*a*) Cross-section.
(*b*) Transformation into *w*-plane.
(*c*) Transformation into *z*-plane.

it was shown that the cross-capacitance per unit length between electrodes $\alpha\beta$ and $\gamma\delta$ is given, when the dielectric thickness is small, by

$$C = \frac{1}{4\pi^2}\log_\epsilon 2 + \frac{3}{16\pi^2}\left(\frac{t}{a}\right)^2\left(1 - \frac{1}{\epsilon^2}\right) + \text{terms in } \left(\frac{t}{a}\right)^2 \text{ and higher} \quad (10)$$

If we transform this system [which is to be compared with that of Fig. 1(*d*)(ii)] back into a circular cross-section by the methods of the previous Section, we obtain the system of Fig. 2(*b*), where the dielectric is still symmetrically distributed with respect to the capacitor symmetry plane but is no longer of constant thickness.

From eqn. (10) it can be seen that for this system each cross-capacitance per unit length is increased by an amount proportional to the square of the dielectric thickness when this thickness is small.

Now consider the electrically symmetrical system of Fig. 3(*a*), which may be compared with Fig. 1(*e*)(i) (with $a = b$). Transforming the system back by the methods of Section 2, we obtain the physically symmetrical systems of Fig. 3(*b*) and (*c*), respectively. It will be seen that the dielectric in the transformed systems is neither symmetrically disposed with respect to the capacitor symmetry plane nor of constant thickness.

A simple calculation on the original system of Fig. 3(*a*) shows that, in this case, the cross-capacitance per unit length between $\alpha\beta$ and $\gamma\delta$ is

$$C = \left(\frac{1}{4\pi^2}\log_\epsilon 2\right)\left[1 + 2\frac{t}{\pi}\frac{\epsilon - 1}{\epsilon} + 4\left(\frac{t}{\pi}\right)^2\left(\frac{\epsilon - 1}{\epsilon}\right)^2 + \ldots\right] \quad (11)$$

i.e. the cross-capacitance per unit length between $\alpha\beta$ and $\gamma\delta$ i increased by an amount directly proportional to the dielectri thickness when the thickness is small.

These results suggest that, for small dielectric thickness, th cross-capacitance per unit length of a symmetric cylindrica capacitor may be written as a Taylor series in terms of th dielectric thickness, in which the coefficient of the linear term i zero, provided that the dielectric is disposed symmetrically wit respect to the capacitor symmetry plane. That this is so wi be demonstrated in Section 5.

Up to this point we have been concerned only with th individual cross-capacitances per unit length. It would be o considerable interest to know what happens to the mean capaci tance per unit length of a symmetrical cylindrical capacitor whe its electrodes have a thin dielectric film distributed on then asymmetrically. Thus in the system of Fig. 3 it may be con jectured that the cross-capacitance per unit length between α and $\beta\gamma$ will be reduced by the same amount, to the first order as the increase in the cross-capacitance per unit length betwee $\alpha\beta$ and $\gamma\delta$. As will be seen in the next Section, such a result i certainly true for the system discussed there (a slightly simplifie version of the system of Fig. 3), and in Section 5 it is shown tha this result is true in general.

(4) THE PARALLEL-PLATE CYLINDRICAL CAPACITOR WIT A UNIFORM DIELECTRIC FILM ON ONE ELECTRODE

By the methods of Section 2 any cylindrical capacitor can b transformed into an equivalent parallel-plate cylindrical capaci tor. When the original capacitor cross-section may be trans formed into one having symmetry about a line passing through a pair of non-adjacent electrode gaps, the dimensions of th equivalent parallel-plate capacitor have been shown to be such

at each cross-capacitance per unit length is given by $C = /4\pi^2) \log_\varepsilon 2$.

In this Section we analyse the general parallel-plate cylindrical pacitor with the additional complication of a dielectric film constant thickness on one electrode. The capacitor cross-ction, together with the co-ordinate system that will be used in e analysis, is shown in Fig. 4.

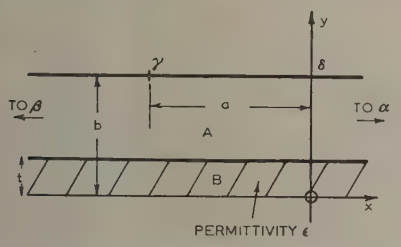


Fig. 4.—Cross-section of parallel-plate cylindrical capacitor with dielectric film on one electrode only.

The working of this Section will be given in some detail, nce the system of Fig. 4 appears to be one of the few general ymmetric cases in which both cross-capacitances per unit ngth can be fairly readily evaluated analytically.

4.1) Cross-Capacitance per Unit Length from $\alpha\beta$ to $\gamma\delta = C_1(\epsilon)$

Electrode $\alpha\beta$ may be placed at unit potential, all other elec-odes at zero. The lines of force are parallel to the y-axis id the dielectric boundary is an equipotential. Thus the quired capacitance per unit length may be regarded as arising om two capacitors in series:

$$[C_1(\epsilon)]^{-1} = \left[\frac{a}{4\pi(b-t)}\right]^{-1} + \left(\frac{\epsilon a}{4\pi t}\right)^{-1} \quad . \quad . \quad (12)$$

id hence

$$C_1(\epsilon) = \frac{a}{4\pi b}\left[1 - \frac{t}{b}\frac{(\epsilon-1)}{\epsilon}\right]^{-1} \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

$$= \frac{a}{4\pi b}\left[1 + \frac{t}{b}\frac{(\epsilon-1)}{\epsilon} + \left(\frac{t}{b}\right)^2\left(\frac{\epsilon-1}{\epsilon}\right)^2 + \ldots\right] \quad (14)$$

In particular, for the symmetrical cylindrical capacitor,

$$\left.\begin{aligned} a &= \log_\varepsilon 2 \\ b &= \pi \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

id substituting eqn. (15) in eqn. (14) and denoting the particular lue of C_1 by C_1^*,

$$C_1^*(\epsilon) = \left(\frac{1}{4\pi^2}\log_\varepsilon 2\right)\left[1 + \frac{t}{\pi}\frac{(\epsilon-1)}{\epsilon} + \left(\frac{t}{\pi}\right)^2\left(\frac{\epsilon-1}{\epsilon}\right)^2 + \ldots\right]$$

. . . . (16)

.2) Cross-Capacitance per Unit Length from $\alpha\delta$ to $\beta\gamma = C_2(\epsilon)$

The calculations here are considerably more difficult. It is onvenient to take electrode $\alpha\delta$ at unit potential and all other ectrodes at zero. In region B, with the co-ordinate system own, the particular solution of Laplace's equation which anishes when $y = 0$ is

$$V_B(x, y) = \int_{-\infty}^{+\infty} B(k)\varepsilon^{jkx}\sinh ky\,dk \quad . \quad . \quad . \quad (17)$$

id, in region A, is

$$V_A(x, y) = \int_{-\infty}^{+\infty} [A_1(k)\cosh ky + A_2(k)\sinh ky]\varepsilon^{jkx}dk \quad (18)$$

with the boundary conditions

$$V_A(x, t) = V_B(x, t) \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

$$\frac{\partial V_A(x, y)}{\partial y} = \epsilon\left[\frac{\partial V_B(x, y)}{\partial y}\right]_{y=t} \quad . \quad . \quad . \quad . \quad (20)$$

These two boundary conditions, together with Lerch's theorem on the uniqueness of Fourier transforms, serve to determine the two functions $A_1(k)$, $A_2(k)$ in terms of $B(k)$:

$$A_1(k) = \tfrac{1}{2}B(k)(1-\epsilon)\sinh 2kt \quad . \quad . \quad . \quad . \quad . \quad (21)$$

$$A_2(k) = \tfrac{1}{2}B(k)[(1+\epsilon) - (1-\epsilon)\cosh 2kt] \quad . \quad (22)$$

On substituting eqns. (21) and (22) in eqn. (18),

$$V_A(x, y) = \tfrac{1}{2}\int_{-\infty}^{+\infty}\{(1-\epsilon)\sinh 2kt\cosh ky + [(1+\epsilon) - (1-\epsilon)\cosh 2kt]\}\sinh kyB(k)\varepsilon^{jkx}dk \quad (23)$$

Setting $y = b$ and using the boundary condition

$$\left.\begin{aligned} V_A(x, b) &= 1 \quad x > 0 \\ &= 0 \quad x < 0 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (24)$$

we obtain, on inverting the Fourier transform of eqn. (18),

$$\tfrac{1}{2}B(k)\{(1-\epsilon)\sinh 2kt\cosh kb + [(1+\epsilon) - (1-\epsilon)\cosh 2kt]\sinh kb\} = \frac{1}{2\pi}\int_0^\infty \varepsilon^{-jkx}dx \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

$$= \frac{1}{2\pi}\left[\frac{1}{jk} + \pi\delta(k)\right] \quad . \quad . \quad . \quad . \quad (26)$$

where $\delta(k)$ is the unit impulse or Dirac delta function.

Using eqn. (26) in eqn. (23) and writing

$$m = \frac{1+\epsilon}{1-\epsilon} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

we find the complete solution for the potential function in region A, namely

$$V_A(x, y) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\frac{1}{jk} + \pi\delta(k)\right]\varepsilon^{jkx} \times \times \frac{\sinh 2kt\cosh ky + (m - \cosh 2kt)\sinh ky}{\sinh 2kt\cosh kb + (m - \cosh 2kt)\sinh kb}dk \quad (28)$$

Now, the charge density on the upper electrode is given by

$$\sigma(x) = -\frac{1}{4\pi}\left[\frac{\partial V_A(x, y)}{\partial y}\right]_{y=b} \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

$$= -\frac{1}{8\pi^2}\int_{-\infty}^{+\infty}\left[\frac{1}{jk} + \pi\delta(k)\right]k\varepsilon^{jkx}F[k, b, t]dk \quad (30)$$

where

$$F[k, b, t] = \frac{\sinh 2kt\sinh kb + (m - \cosh 2kt)\cosh kb}{\sinh 2kt\cosh kb + (m - \cosh 2kt)\sinh kb} \quad (31)$$

It seems impossible to evaluate eqn. (30) in closed form, so we content ourselves with an expansion in powers of t, the dielectric thickness. This procedure is reasonable as we are primarily concerned with the effects of thin dielectric films. Thus,

expanding eqn. (31) in powers of t, and substituting the resulting expansion in eqn. (30),

$$\sigma(x) = -\frac{1}{8\pi^2}\int_{-\infty}^{+\infty}\left[\frac{1}{jk}+\pi\delta(k)\right]k\varepsilon^{jkx}\frac{\cosh kb}{\sinh kb}dk$$
$$-t\left(\frac{2}{1-m}\right)\frac{1}{8\pi^2}\int_{-\infty}^{+\infty}\left[\frac{1}{jk}+\pi\delta(k)\right]k^2\varepsilon^{jkx}\frac{1}{\sinh^2 kb}dk$$
$$-t^2\left(\frac{2}{1-m}\right)^2\frac{1}{8\pi^2}\int_{-\infty}^{+\infty}\left[\frac{1}{jk}+\pi\delta(k)\right]k^3\varepsilon^{jkx}\frac{\cosh kb}{\sinh^3 kb}dk$$
$$+\text{ terms in } t^3 \text{ and higher} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (32)$$

The three integrals occurring here are readily evaluated, for negative x, by contour integration. Results are given in Appendix 8.1. On using these results in eqn. (32) we have, finally,

$$\sigma(x)=\frac{1}{4\pi b}\left\{\frac{\varepsilon^{\pi x/b}}{(1-\varepsilon^{\pi x/b})}+\frac{t}{b}\frac{(\epsilon-1)}{\epsilon}\left[\frac{\varepsilon^{\pi x/b}}{1-\varepsilon^{\pi x/b}}+\frac{\frac{\pi x}{b}\varepsilon^{\pi x/b}}{(1-\varepsilon^{\pi x/b})^2}\right]\right.$$
$$+\frac{1}{2}\left(\frac{t}{b}\right)^2\left(\frac{\varepsilon-1}{\epsilon}\right)^2\left[\frac{2\varepsilon^{\pi x/b}}{1-\varepsilon^{\pi x/b}}+\frac{4\frac{\pi x}{b}\varepsilon^{\pi x/b}}{(1-\varepsilon^{\pi x/b})^2}\right.$$
$$\left.\left.+\frac{\left(\frac{\pi x}{b}\right)^2(1+\varepsilon^{\pi x/b})\varepsilon^{\pi x/b}}{(1-\varepsilon^{\pi x/b})^3}\right]\right\}$$
$$+\text{ terms in } (t/b)^3 \text{ and higher} \quad (33)$$

Here we have reintroduced ϵ by the relation

$$\frac{2}{1-m}=\frac{\epsilon-1}{\epsilon} \quad . \quad . \quad . \quad . \quad . \quad . \quad (34)$$

The cross-capacitance per unit length, $C_2(\epsilon)$, is just the total charge on electrode $\beta\gamma$, namely

$$C_2(\epsilon)=\int_{-\infty}^{-a}\sigma(x)dx \quad . \quad . \quad . \quad . \quad . \quad . \quad (35)$$

We now substitute eqn. (33) in eqn. (35), and carry out the integration, making use of the results given in Appendix 8.2.

$$C_2(\epsilon) = -\frac{1}{4\pi^2}\log_\varepsilon(1-\varepsilon^{-\pi a/b})$$
$$-\frac{1}{4\pi^2}\left(\frac{t}{b}\right)\left(\frac{\epsilon-1}{\epsilon}\right)\frac{\pi a}{b}\frac{\varepsilon^{-\pi a/b}}{1-\varepsilon^{-\pi a/b}}$$
$$+\frac{1}{4\pi^2}\left(\frac{t}{b}\right)^2\left(\frac{\epsilon-1}{\epsilon}\right)^2\frac{1}{2}\left[\left(\frac{\pi a}{b}\right)^2\frac{\varepsilon^{-\pi a/b}}{(1-\varepsilon^{-\pi a/b})^2}-2\frac{\pi a}{b}\frac{\varepsilon^{-\pi a/b}}{1-\varepsilon^{-\pi a/b}}\right]$$
$$+\text{ terms in } \left(\frac{t}{b}\right)^3 \text{ and higher} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)$$

Of particular interest is the case of the symmetrical cylindrical capacitor for which

$$a=\log_\varepsilon 2 \qquad b=\pi \quad . \quad . \quad . \quad . \quad . \quad (37)$$

For this case, denoting the particular value of $C_2(\epsilon)$ by C^*,

$$C_2^*(\epsilon)=\frac{1}{4\pi^2}\log_\varepsilon 2\left[1-\frac{t}{\pi}\frac{\epsilon-1}{\epsilon}-\left(\frac{t}{\pi}\right)^2\left(\frac{\epsilon-1}{\epsilon}\right)^2(1-\log_\varepsilon 2)\right.$$
$$\left.+\text{ terms in } \left(\frac{t}{\pi}\right)^3 \text{ and higher}\right] \quad (38)$$

We note that the linear terms in eqns. (16) and (38) are equ in magnitude and opposite in sign, so that the change in t mean capacitance per unit length is only of the second order the dielectric thickness. Specifically

$$\frac{C_1^*(\epsilon)+C_2^*(\epsilon)}{2}=\frac{1}{4\pi^2}\log_\varepsilon 2\left[1+\left(\frac{t}{\pi}\right)^2\left(\frac{\epsilon-1}{\epsilon}\right)^2\frac{\log_\varepsilon 2}{2}\right.$$
$$\left.+\text{ terms in } \left(\frac{t}{\pi}\right)^3 \text{ and higher}\right] \quad (3$$

This example thus supports the conjecture that, in the ca of a thin dielectric film asymmetrically distributed on the ele trodes of a cylindrical capacitor, the mean capacitance per un length remains constant to the first order.

Finally let us briefly consider the results for the gener parallel-plate cylindrical capacitor with no dielectric fil Setting $t = 0$ in eqns. (14) and (36) we have (on dropping t arguments of C_1 and C_2)

$$C_1=\frac{a}{4\pi b} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4$$

$$C_2=-\frac{1}{4\pi^2}\log_\varepsilon(1-\varepsilon^{-\pi a/b}) \quad . \quad . \quad . \quad (4$$

As a partial check on the fairly complicated calculations this Section we note that the introduction of a parameter

$$\tau=\varepsilon^{\pi a/b}-1 \quad . \quad . \quad . \quad . \quad . \quad (4$$

enables eqns. (40) and (41) to be written in the standard par metric form of eqns. (4) and (5).

(5) GENERAL RESULTS CONCERNING DIELECTRIC FILMS IN CYLINDRICAL CAPACITORS

In this Section we give proofs of the following general resul which have been suggested by the working of Sections 3 and

(*a*) In an electrically symmetric cylindrical capacitor, arbitrary thin dielectric film on the capacitor electrodes produc only a second-order change in the mean capacitance per un length.

(*b*) In a physically symmetric cylindrical capacitor, a th dielectric film on the capacitor electrodes, which is dispos symmetrically with respect to the capacitor symmetry plan produces only a second-order change in each cross-capacitan per unit length.

Before proceeding to the proof of these results we establi a preliminary lemma. This lemma states that, given a cylindrical capacitor with a thin dielectric film on its electrod there exists an equivalent cylindrical capacitor containing dielectric,† whose electrodes lie wholly within the space occupi by the dielectric film, such that the electric fields, at all poi in the dielectricless region of the two capacitors, are equal the first order.

The significance of this lemma is that it shows, for dielect films whose thickness is small, that the perturbation required obtain the equivalent electrode system is both small, of t same order as the dielectric thickness, being in fact bounded the dielectric itself, and independent of the field configuratio i.e. independent of the individual electrode potentials.

Thus consider the cross-section of an arbitrary cylindri capacitor and, in particular, a small region near electrode o say (assumed at zero potential without loss of generality), shown in Fig. 5.

† Throughout the paper the term dielectric is used in the sense of a mate substance.

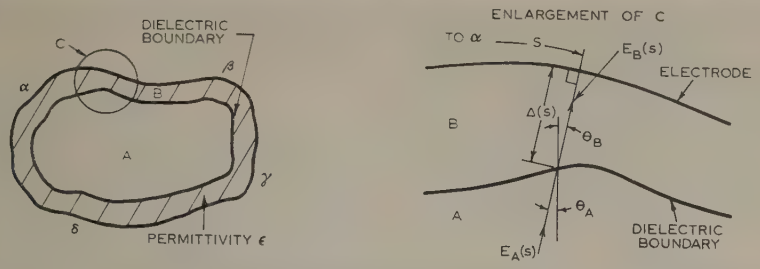


Fig. 5.—Cross-section of general cylindrical capacitor with arbitrary thin dielectric film.

In the following discussion we assume that:

(i) $\Delta(s)$ is small compared with the length of the line of force rminating at s.

(ii) $\Delta'(s)$ is such that the field is almost normal to the dielectric oundary.

With these assumptions, the field $E_B(s)$ in the dielectric can regarded as essentially constant for a given s and, on equating rmal components of electric displacement,

$$E_A(s)\cos\theta_A = \epsilon E_B(s)\cos\theta_B \quad . \quad . \quad . \quad . \quad . \quad (43)$$

$$E_B(s) = \frac{1}{\epsilon}E_A(s)\sqrt{\frac{1+\tan^2\theta_B}{1+\tan^2\theta_A}} \quad . \quad . \quad . \quad (44)$$

sing the identity

$$\epsilon\tan\theta_A = \tan\theta_B \quad . \quad . \quad . \quad . \quad . \quad . \quad (45)$$

nd noting that for small $\Delta'(s)$

$$\tan\theta_B = \Delta'(s) + o[\Delta^2(s)] \quad . \quad . \quad . \quad . \quad (46)$$

e have

$$E_B(s) = \frac{1}{\epsilon}E_A(s)\left\{\sqrt{\frac{1+[\Delta'(s)]^2}{1+\left[\frac{\Delta'(s)}{\epsilon}\right]^2}} + o[\Delta^2(s)]\right\} \quad (47)$$

$$= \frac{1}{\epsilon}E_A(s)\{1 + o[\Delta^2(s)]\} \quad . \quad . \quad . \quad . \quad . \quad (48)$$

hus for small $\Delta(s)$ the potential function at the dielectric oundary is

$$V(s) = \frac{1}{\epsilon}\Delta(s)E_A(s) + o[\Delta^2(s)] \quad . \quad . \quad . \quad (49)$$

Now consider the same region of the capacitor but with the ielectric removed and the electrode surface at s advanced a istance $x\Delta(s)$ towards the former position of the dielectric oundary. The potential at the dielectric boundary is now mply

$$V^*(s) = (1-x)\Delta(s)E_A^*(s) + o[\Delta^2(s)] \quad . \quad . \quad (50)$$

here asterisks have been used to denote the values of $V(s)$ and $_A(s)$ in this situation.

Clearly, if $V(s) = V^*(s)$ for all s, then so must the electric elds in region A also be equal; in particular we must have $_A(s) = E_A^*(s)$ for all s. Equating the two expressions for $V(s)$ nd solving for x,

$$x = \frac{\epsilon - 1}{\epsilon} + o[\Delta(s)] \quad . \quad . \quad . \quad . \quad (51)$$

hich result is independent of s to the first order and also ndependent of the way in which the field arises.

Thus we have shown that the electric field in the region noccupied by dielectric remains constant to the first order, if ne dielectric and electrode systems are replaced by a perturbed ectrode system (without dielectric) such that each new electrode ement is a distance $\Delta(s)(\epsilon - 1)/\epsilon$ from the corresponding original electrode position. Since $(\epsilon - 1)/\epsilon < 1$, the perturbed electrode system of the equivalent dielectricless capacitor lies wholly within the region formerly occupied by the dielectric, which proves the lemma.

Finally, the capacitance per unit length between electrodes $\alpha\beta$ and $\gamma\delta$, say, can be computed by assuming that electrode $\gamma\delta$ is at unit potential and all other electrodes are at zero potential. Hence

$$C_1 = \int_\alpha^\beta \sigma(s)ds \quad . \quad . \quad . \quad . \quad . \quad (52)$$

where $\sigma(s)$, the charge density per unit length, may be written in terms of the electric displacement at the electrode surface. To the first order, this electric displacement is identical with the normal component of the electric field at the dielectric, and hence

$$\sigma(s) = \frac{1}{4\pi}E_A(s)\left\{1 + \left[\frac{\Delta'(s)}{\epsilon}\right]^2\right\}^{-1/2} + o[\Delta(s)] \quad (53)$$

$$= \frac{1}{4\pi}E_A(s) + o[\Delta(s)] \quad . \quad . \quad . \quad . \quad . \quad . \quad (54)$$

We have already seen that the electric field $E_A(s)$ remains constant to the first order in the equivalent dielectricless capacitor, and so, on using eqn. (54) in eqn. (52), it is clear that the same conclusion applies to the capacitance per unit length, C_1, and by similar working to the other cross-capacitance per unit length, C_2.

As an example of the use of this lemma we may compute the first-order changes in the cross-capacitances per unit length for the system of Fig. 4 due to the presence of the dielectric film. Then, according to the lemma, the equivalent dielectricless capacitor would have the same disposition, but the distance b between the electrodes must now be reduced by an amount $t(\epsilon - 1)/\epsilon$. Thus, using eqns. (40) and (41),

$$C_1(\epsilon) = \frac{a}{4\pi\left(b - t\frac{\epsilon-1}{\epsilon}\right)} \text{ to the first order} \quad . \quad . \quad . \quad (55)$$

$$C_2(\epsilon) = -\frac{1}{4\pi^2}\log_\varepsilon\left[1 - \exp\left(-\frac{\pi a}{b - t\frac{\epsilon-1}{\epsilon}}\right)\right]$$

to the first order (56)

On expanding eqns. (55) and (56) in powers of $t(\epsilon - 1)/\epsilon$,

$$C_1(\epsilon) = \frac{a}{4\pi b}\left(1 + \frac{t}{b}\frac{\epsilon-1}{\epsilon} + \text{higher terms}\right) \quad . \quad . \quad (57)$$

$$C_2(\epsilon) = -\frac{1}{4\pi^2}\log_\varepsilon(1 - \varepsilon^{-\pi a/b})$$

$$-\frac{1}{4\pi^2}\frac{t}{b}\frac{\epsilon-1}{\epsilon}\frac{\pi a}{b}\frac{\varepsilon^{-\pi a/b}}{1-\varepsilon^{-\pi a/b}} + \text{higher terms} \quad (58)$$

which of course agree with eqns. (14) and (36), to the first order.

To prove the first general result, mentioned at the beginning of this Section, consider an electrically symmetric cylindrical capacitor with an arbitrary thin dielectric film on all its electrodes. By the lemma and its subsequent discussion it follows that there is an equivalent dielectricless capacitor, with the same cross-capacitances per unit length to the first order, but whose electrodes lie wholly within the region formerly occupied by dielectric. The change in the mean capacitance per unit length, C_M, is then given by

$$\delta C_M = \delta\left(\frac{C_1 + C_2}{2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (59)$$

$$= \tfrac{1}{2}(dC_1 + dC_2) + \text{terms of higher degree} \quad . \quad (60)$$

But for the electrically symmetric cylindrical capacitor ($C_1 = C_2$) we have, from eqn. (9),

$$dC_1 = -\,dC_2 \text{ (if } C_1 = C_2) \quad . \quad . \quad . \quad . \quad (61)$$

and so

$$\delta C_M = 0 \text{ (to the first order)} \quad . \quad . \quad . \quad (62)$$

which proves the first result.

To prove the second general result, consider a physically symmetric cylindrical capacitor C with a thin dielectric film on all its electrodes, which is disposed symmetrically with respect to the capacitor symmetry plane. The dielectric in such a capacitor may be considered to be introduced in two steps as shown in Fig. 6.

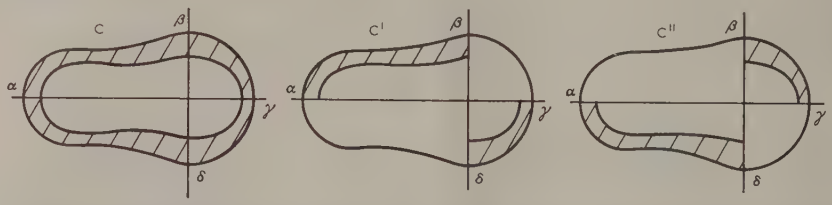


Fig. 6.—Sequential placement of dielectric film in a physically symmetric cylindrical capacitor.

When we consider the equivalent dielectricless capacitors, which exist by the lemma above, we have, on making use of eqn. (9),

$$\delta C_1' = -\,\delta C_2' \text{ (to first order)} \quad . \quad . \quad . \quad . \quad (63)$$

$$\delta C_1'' = -\,\delta C_2'' \text{ (to first order)} \quad . \quad . \quad . \quad . \quad (64)$$

Also, because of the symmetry of the dielectric film,

$$\delta C_1' = \delta C_2'' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (65)$$

$$\delta C_2' = \delta C_1'' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (66)$$

On superposing the two sets of equations to obtain results for the whole capacitance, C,

$$\delta C_1 = \delta C_1' + \delta C_1'' = 0 \text{ (to first order)} \quad . \quad . \quad (67)$$

$$\delta C_2 = \delta C_2' + \delta C_2'' = 0 \text{ (to first order)} \quad . \quad . \quad (68)$$

which proves the second general result, namely that in this case the individual cross-capacitances per unit length remain constant to the first order.

Alternatively we may note that, by the lemma, the equivalent dielectricless capacitor is also physically symmetrical, and hence to the first order the individual cross-capacitances per unit length are both $(1/4\pi^2)\log_e 2$ and thus remain constant to the first order which proves the second general result.

(6) CONCLUSION

The application of conformal-transformation methods to certain types of cylindrical 3-terminal capacitors, whose electrodes carry thin dielectric films, has suggested some general results concerning the effects on the cross-capacitances of such dielectric films. The detailed computation of the cross-capacitances of a parallel-plate cylindrical capacitor (with a thin highly asymmetric dielectric film on its electrodes) has further substantiated these conjectures, and proofs are given in the final Section.

It should be realized that the results of the paper apply only to dielectric films which maintain a constant cross-section throughout the length of the capacitor. More generally, dielectric films will be distributed in a 3-dimensional pattern over the electrode surfaces. In this case the potential problem is correspondingly 3-dimensional and hence the powerful methods of conformal transformation are no longer available for its solution. Such problems require further study.

(7) REFERENCES

(1) THOMPSON, A. M., and LAMPARD, D. G.: 'A New Theorem in Electrostatics and its Application to Calculable Standards of Capacitance', *Nature*, 1956, **177**, p. 888.

(2) LAMPARD, D. G.: 'A New Theorem in Electrostatics with Applications to Calculable Standards of Capacitance', *Proceedings I.E.E.*, Monograph No. 216 M, January, 1957 (**104** C, p. 271).

(3) COPSON, E. T.: 'Theory of Functions of a Complex Variable' (Oxford, 1948).

(4) LAMPARD, D. G.: C.S.I.R.O. Division of Electrotechnology, Internal Report, March, 1957.

(5) VAN DER PAUW, L. J.: 'A Method of Measuring Specific Resistivity and Hall Effects of Discs of Arbitrary Shape', *Philips Research Reports*, 1958, **13**, p. 1.

(6) THOMPSON, A. M.: 'The Cylindrical Cross-Capacitor as a Calculable Standard', *Proceedings I.E.E.*, Paper No. 2887 M, May, 1959 (**106** B, p. 307).

(8) APPENDIX

(8.1) Values of Integrals occurring in Eqn. (32)

The integrals occurring in eqn. (32) may be evaluated (for negative x) by summing residues at the poles $k = -jn\pi/b$ with $n = 0, 1, 2 \ldots$ The results are:

$$\frac{1}{8\pi^2}\int_{-\infty}^{+\infty}\left[\frac{1}{jk} + \pi\delta(k)\right]k\varepsilon^{jkx}\frac{\cosh kb}{\sinh kb}dk = -\frac{1}{4\pi b}\sum_{n=1}^{\infty}\varepsilon^{\pi nx/b} \quad (69)$$

$$= -\frac{1}{4\pi b}\frac{\varepsilon^{\pi x/b}}{1-\varepsilon^{\pi x/b}} \quad (70)$$

$$\int_{-\infty}^{+\infty}\left[\frac{1}{jk}+\pi\delta(k)\right]k^2\varepsilon^{jkx}\frac{1}{\sinh^2 kb}dk$$

$$=-\frac{1}{4\pi b^2}\sum_{n=1}^{\infty}\left(1+\frac{\pi nx}{b}\right)\varepsilon^{\pi nx/b} \quad . \quad (71)$$

$$=-\frac{1}{4\pi b^2}\left[\frac{1-\varepsilon^{\pi x/b}}{\varepsilon^{\pi x/b}}+\frac{\frac{\pi x}{b}\varepsilon^{\pi x/b}}{(1-\varepsilon^{\pi x/b})^2}\right] \quad (72)$$

$$\int_{-\infty}^{+\infty}\left[\frac{1}{jk}+\pi\delta(k)\right]k^3\varepsilon^{jkx}\frac{\cosh kb}{\sinh^3 kb}dk$$

$$=-\frac{1}{8\pi b^3}\sum_{n=1}^{\infty}\left[\left(\frac{\pi nx}{b}\right)^2+4\left(\frac{\pi nx}{b}\right)+2\right]\varepsilon^{\pi nx/b} \quad (73)$$

$$=-\frac{1}{8\pi b^3}\left[2\frac{\varepsilon^{\pi x/b}}{1-\varepsilon^{\pi x/b}}+4\frac{\frac{\pi x}{b}\varepsilon^{\pi x/b}}{(1-\varepsilon^{\pi x/b})^2}+\frac{\left(\frac{\pi x}{b}\right)^2(1+\varepsilon^{\pi x/b})\varepsilon^{\pi x/b}}{(1-\varepsilon^{\pi x/b})^3}\right] \quad (74)$$

(8.2) Values of Integrals occurring in Eqn. (35)

When carrying out the integration of eqn. (35), with $\sigma(x)$ given by eqn. (33), the following results are needed:

$$\int_{-\infty}^{-a}\frac{\varepsilon^{\pi x/b}}{1-\varepsilon^{\pi x/b}}dx=-\frac{b}{\pi}\log_\varepsilon(1-\varepsilon^{-\pi a/b}) \quad . \quad . \quad . \quad . \quad . \quad . \quad (75)$$

$$\int_{-\infty}^{-a}\frac{\frac{\pi x}{b}\varepsilon^{\pi x/b}}{(1-\varepsilon^{\pi x/b})^2}dx=-\frac{b}{\pi}\left[\frac{\frac{\pi a}{b}\varepsilon^{-\pi a/b}}{1-\varepsilon^{-\pi a/b}}-\log_\varepsilon(1-\varepsilon^{-\pi a/b})\right] \quad (76)$$

$$\int_{-\infty}^{-a}\frac{\left(\frac{\pi x}{b}\right)^2(1+\varepsilon^{\pi x/b})\varepsilon^{\pi x/b}}{(1-\varepsilon^{\pi x/b})^3}dx$$

$$=\frac{b}{\pi}\left[\frac{\left(\frac{\pi a}{b}\right)^2\varepsilon^{-\pi a/b}}{(1-\varepsilon^{-\pi a/b})^2}+\frac{2\frac{\pi a}{b}\varepsilon^{-\pi a/b}}{(1-\varepsilon^{-\pi a/b})}-2\log_\varepsilon(1-\varepsilon^{-\pi a/b})\right] \quad (77)$$

621.376.3

The Institution of Electrical Engineers
Monograph No. 352
Jan. 19

EXPLICIT FORM OF F.M. DISTORTION PRODUCTS WITH WHITE-NOISE MODULATION

By R. G. MEDHURST, B.Sc.

(*The paper was first received 27th July, and in revised form 10th September*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in January*, 1960.)

SUMMARY

When a frequency-modulated wave passes through a network whose phase or amplitude characteristics vary non-linearly with frequency, distortion terms appear as both frequency and amplitude modulation of the output wave. If the characteristics are expressed as power series, these distortion terms appear, to first order, as products of powers of time derivatives of the unwanted frequency modulation. When the frequency modulation may be simulated by a band of random noise (as in multiplex telephony carrying large numbers of channels), the spectra of the distortion products can, in principle, be described by simple algebraic functions of the characteristics (i.e. the minimum and maximum frequencies and the r.m.s. frequency deviation) of the modulating noise band.

Except in certain special cases, the derivation of these algebraic formulae by straightforward analytical methods becomes prohibitively tedious for distortions of order much above the second. However, once the formulae are found, the insertion of numerical values for particular cases is straightforward. In the present paper it is shown how the problem can be reduced to the repetition of a number of standard operations which can be carried out using a digital computer. The technique is illustrated by application to fourth-order distortion appearing in the amplitude modulation, generated by terms in the amplitude characteristic up to sixth degree. Even in such an apparently simple case as this it appears from the literature that the closed form of the distortion formula has not hitherto been obtained. This example is of direct practical interest since, for example, the amplitude characteristic of a maximally-flat-amplitude triple-tuned circuit is of sixth degree in the region around the mid-band frequency. With a minor modification, the resulting formula also applies to fourth-order distortion appearing in the frequency modulation, owing to terms in the phase characteristic up to sixth degree.

Use is made of a discontinuous contour integral applied to a similar, but somewhat simpler, case by Bennett;[1] a closely analogous course can be followed using the more recently developed theory of generalized functions,[2] but in this particular problem the contour-integral method is more economical.

Formulae for the various orders of distortion in the top channel due to amplifier and discriminator characteristics are given in tabular form.

LIST OF SYMBOLS

ω_c = Carrier frequency, rad/sec.
ω_M = Undistorted frequency modulation, rad/sec.
ω_Δ = R.M.S. frequency deviation, rad/sec.
μ_t = Phase modulation, rad.
$\hat{\omega}_m$, $\check{\omega}_m$ = Maximum and minimum modulating frequency, rad/sec.
ω_m = A representative frequency of the noise-band modulation, rad/sec.
$\phi_{\omega m}$ = A random phase angle associated with the tone of frequency ω_m, rad.
ω = Difference between a general angular frequency and ω_c, rad/sec.
a_n, b_n = Respective coefficients of nth-order terms in expansions of amplifier phase and amplitude characteristics, in powers of ω/ω_c.
a_n', b_n' = Respective coefficients of nth-order terms in expansions of discriminator phase and amplitude characteristics, in powers of ω/ω_c.
B_0, B_1 = Respective coefficients of zero and first-order terms in expansion of discriminator amplitude characteristic, in powers of ω/ω_c.
$c_n' = B_0 a_n' + B_1 a_{n-1}'$.
$A = \frac{b_4'}{b_6'}\omega_c^2 + 15\omega_\Delta^2$.
$a = \frac{\omega_\Delta\sqrt{2}}{\sqrt{(\hat{\omega}_m - \check{\omega}_m)}}$.

(1) INTRODUCTION

A problem which constantly recurs in the design and production of f.m. trunk radio systems carrying many f.d.m. speech channels is the minimization of intermodulation distortion. Distortion of this kind may be generated by echoes in the radio transmission path or in the feeders, by interference of other carriers, by non-linear modulator characteristics or by non-linearities in the amplitude and phase characteristics of amplifiers and discriminators. The first two distortion mechanisms may now be considered to be reasonably well understood,[3–8] but the problem of the effect of non-linear transmission characteristics has so far been made amenable to analysis only when the modulation takes the form of a single tone,[9–11] or for special forms of transmission characteristic.[12–14]

When the amplitude and phase characteristics of amplifiers or discriminators are expressed as power series in terms of the frequency deviation it can be shown[9, 15] that distortion terms appearing both as frequency modulation and amplitude modulation of the output wave appear as products of powers of time derivatives of the frequency modulation; this result holds if the departures from linearity of the characteristics are not too great.[13] Thus one has to deal with expressions of the form

$$k\left[\frac{d^l\omega_M}{dt^l}\right]^p\left[\frac{d^m\omega_M}{dt^m}\right]^q\left[\frac{d^n\omega_M}{dt^n}\right]^r \quad . \quad . \quad . \quad (1$$

When large numbers of speech channels are involved, the f.d.m. signal (in which the speech channels occupy 4 kc/s bands spaced in frequency one above another) can be satisfactorily simulated by a band of white noise.[16, 17] Thus, if this signal is used directly to frequency-modulate a carrier, the frequency modulation can be written in the form

$$\omega_M = a\sum_{\omega=\check{\omega}_m}^{\hat{\omega}_m}\cos(\omega_m t + \phi_{\omega m}) \quad . \quad . \quad . \quad . \quad (2$$

where ω_m increases in unit steps.

If the right-hand side of eqn. (2) is inserted into expression (1) there can be extracted, when the total order of the term is odd

Correspondence on Monographs is invited for consideration with a view to publication.
The paper is a communication from the Staff of the Research Laboratories of The General Electric Company Limited, Wembley, England.

portion coherent with μ_t (though possibly not with the same ectral distribution). This will not produce distortion in a ultiplex-telephony system, although it may distort a television gnal. The remainder will consist of higher-order portions, ther coherent or incoherent with terms of lower order. Both pes of term contribute distortion, which appears as unintel-gible crosstalk, and they have to be considered separately, since ey add in different ways to the appropriate lower-order terms. One is interested in practice in the magnitude of the distortion ppearing in each channel, relative to the wanted speech power. onsidering the channels as occupying bands which are very arrow compared with the total base band, it will be sufficient evaluate the spectra of the distortion terms and to compare e power at particular frequencies with the wanted signal power the same frequencies.

Evaluation of spectra via the auto-correlation function and e Wiener–Khintchine relation has been carried through with arying success in a range of problems concerning modulation ith noise and interference by noise (e.g. calculation of echo istortion,[4–6] of the r.f. spectrum of waves frequency-modulated y noise [8, 18, 19] and of the a.m. and f.m. disturbance of a carrier y added noise[24]). The application of this method to the xpansion in spectral form of sums of terms such as expres-on (1), while straightforward in principle, results in analytical mplexity which becomes very formidable as the order of rms increases.

When ω_M has the form of expression (2) the spectrum of nctions such as expression (1) can be evaluated, in principle, y direct multiplication and summation of all resulting tones of e same frequency, with due regard to coherence among oups of terms. This was done by Lewin[3] in a penetrating nalysis of the spectrum of the nth power of a band of white ise. However, the labour of applying similar methods to the ore complicated distortion products appearing in practice ould be extremely formidable. What is shown in the present aper is that after some preliminary rearrangement the problem n be reduced to repeated application of two simple rules of gebraic manipulation, both of which can be programmed into digital computer.

(2) FORM OF THE DISTORTED OUTPUT FROM AN AMPLIFIER-LIMITER-DISCRIMINATOR CHAIN

The distortion products we wish to evaluate arise from phase d amplitude non-linearities in amplifiers and discriminators. e amplifier amplitude characteristic will normally be approxi-ately constant over the significant r.f. band. Constant and ear terms in the amplifier phase characteristic will merely oduce respectively a phase shift of the carrier and a delay on e intelligence; thus the major portions of such terms can be btracted off initially. Consequently, the amplifier phase and nplitude characteristics can be written respectively as

$$\phi(\omega_c + \omega) = \sum_{n=0}^{\infty} a_n\left(\frac{\omega}{\omega_c}\right)^n \quad . \quad . \quad . \quad . \quad (3)$$

$$1 + \rho(\omega_c + \omega) = 1 + \sum_{n=0}^{\infty} b_n\left(\frac{\omega}{\omega_c}\right)^n \quad . \quad . \quad (4)$$

ere the a's and b's are of such magnitudes that the output stortion is small. The discriminator network will normally ve a substantial linear term, its function being to generate opy of the frequency modulation in the amplitude modulation the outcoming wave. Thus, the discriminator phase and nplitude characteristics will be written as follows (where pital letters are used they are intended as a reminder that the fects of these terms are not small):

$$\phi'(\omega_c + \omega) = \sum_{n=0}^{\infty} a'_n\left(\frac{\omega}{\omega_c}\right)^n \quad . \quad . \quad . \quad . \quad (5)$$

$$\rho'(\omega_c + \omega) = B_0 + B_1\left(\frac{\omega}{\omega_c}\right) + \sum_{n=2}^{\infty} b'_n\left(\frac{\omega}{\omega_c}\right)^n . \quad . \quad (6)$$

In practice, amplitude modulation generated by the amplifier will be largely suppressed by a subsequent limiter. Under these circumstances, it can be shown (Section 6.1) that the output of the discriminator takes the approximate form

$$\begin{aligned}
&B_0 + \frac{B_1}{\omega_c}\omega_M + \frac{B_1}{\omega_c}\Big[2\frac{a_2}{\omega_c^2}\omega_M\omega_M^{\mathrm{I}} + 3\frac{a_3}{\omega_c^3}\omega_M^2\omega_M^{\mathrm{I}} + \frac{a_4}{\omega_c^4}(4\omega_M^3\omega_M^{\mathrm{I}} \\
&\quad - 10\omega_M^{\mathrm{I}}\omega_M^{\mathrm{II}} - 4\omega_M\omega_M^{\mathrm{III}}) + \frac{a_5}{\omega_c^5}(5\omega_M^4\omega_M^{\mathrm{I}} - 50\omega_M\omega_M^{\mathrm{I}}\omega_M^{\mathrm{II}} \\
&\quad - 10\omega_M^2\omega_M^{\mathrm{III}} - 15\omega_M^{\mathrm{I}3}) + \frac{a_6}{\omega_c^6}(6\omega_M^5\omega_M^{\mathrm{I}} - 90\omega_M\omega_M^{\mathrm{I}3} \\
&\quad - 150\omega_M^2\omega_M^{\mathrm{I}}\omega_M^{\mathrm{II}} - 20\omega_M^3\omega_M^{\mathrm{III}} + 21\omega_M^{\mathrm{I}}\omega_M^{\mathrm{IV}} \\
&\quad + 6\omega_M\omega_M^{\mathrm{V}} + 35\omega_M^{\mathrm{II}}\omega_M^{\mathrm{III}}) + \ldots \\
&\quad - 3\frac{b_3}{\omega_c^3}(\omega_M^{\mathrm{I}2} + \omega_M\omega_M^{\mathrm{II}}) - 6\frac{b_4}{\omega_c^4}(2\omega_M\omega_M^{\mathrm{I}2} + \omega_M^2\omega_M^{\mathrm{II}}) \\
&\quad + \frac{b_5}{\omega_c^5}(-30\omega_M^2\omega_M^{\mathrm{I}2} - 10\omega_M^3\omega_M^{\mathrm{II}} + 15\omega_M^{\mathrm{I}}\omega_M^{\mathrm{III}} + 5\omega_M\omega_M^{\mathrm{IV}} \\
&\quad + 10\omega_M^{\mathrm{II}2}) + \frac{b_6}{\omega_c^6}(-60\omega_M^3\omega_M^{\mathrm{I}2} - 15\omega_M^4\omega_M^{\mathrm{II}} + 105\omega_M^{\mathrm{I}2}\omega_M^{\mathrm{II}} \\
&\quad + 60\omega_M\omega_M^{\mathrm{II}2} + 90\omega_M\omega_M^{\mathrm{I}}\omega_M^{\mathrm{III}} + 15\omega_M^2\omega_M^{\mathrm{IV}}) + \ldots\Big] \\
&\quad + \Big[3\frac{c'_3}{\omega_c^3}\omega_M\omega_M^{\mathrm{I}} + 6\frac{c'_4}{\omega_c^4}\omega_M^2\omega_M^{\mathrm{I}} + \frac{c'_5}{\omega_c^5}(10\omega_M^3\omega_M^{\mathrm{I}} - 5\omega_M\omega_M^{\mathrm{III}} \\
&\quad - 10\omega_M^{\mathrm{I}}\omega_M^{\mathrm{II}}) + \frac{c'_6}{\omega_c^6}(15\omega_M^4\omega_M^{\mathrm{I}} - 15\omega_M^{\mathrm{I}3} - 60\omega_M\omega_M^{\mathrm{I}}\omega_M^{\mathrm{II}} \\
&\quad - 15\omega_M^2\omega_M^{\mathrm{III}}) + \ldots + \frac{b'_2}{\omega_c^2}\omega_M^2 + \frac{b'_3}{\omega_c^3}\omega_M^3 + \frac{b'_4}{\omega_c^4}(\omega_M^4 - 3\omega_M^{\mathrm{I}2} \\
&\quad - 4\omega_M\omega_M^{\mathrm{II}}) + \frac{b'_5}{\omega_c^5}(\omega_M^5 - 10\omega_M^2\omega_M^{\mathrm{II}} - 15\omega_M\omega_M^{\mathrm{I}2}) \\
&\quad + \frac{b'_6}{\omega_c^6}(\omega_M^6 - 45\omega_M^2\omega_M^{\mathrm{I}2} - 20\omega_M^3\omega_M^{\mathrm{II}} + 6\omega_M\omega_M^{\mathrm{IV}} \\
&\quad + 15\omega_M^{\mathrm{I}}\omega_M^{\mathrm{III}} + 10\omega_M^{\mathrm{II}2}) + \ldots\Big] \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)
\end{aligned}$$

In this expression, only the distortion terms incoherent with the wanted signal have been retained. The superscripts attached to the ω_M terms denote differentiation with respect to time. Thus ω_M^{II} signifies $(d^2/dt^2)\omega_M$ and $\omega_M^{\mathrm{I}2}$ signifies $(d\omega_M/dt)^2$.

(3) REDUCTION OF EXPLICIT FORM OF FOURTH-ORDER DISTORTION DUE TO TERMS OF THE DISCRIMINATOR AMPLITUDE CHARACTERISTIC UP TO SIXTH POWER

In an earlier paper[15] the explicit form of the third-order distortion associated with the discriminator amplitude characteristic, taking account of terms in this characteristic up to the fifth power, was evaluated by direct integration. Consideration of the details of this analysis will quickly make it plain that no more complex case is likely to be attempted by the same method. The range of the technique advocated in the present paper is, however, limited only by the availability of sufficient digital-computer

running time. The method will best be illustrated by an example, which will be taken as the reduction to algebraic form of the fourth-order distortion due to terms of the discriminator amplitude characteristic up to sixth power. With a minor modification, the result can also be made to give fourth-order distortion due to corresponding terms of the amplifier phase characteristic, as explained in Section 4.

The terms in expression (7) associated with the discriminator amplitude characteristic which contribute to fourth-order distortion are

$$\frac{b_4'}{\omega_c^4}\omega_M^4 + \frac{b_6'}{\omega_c^6}(\omega_M^6 - 45\omega_M^2\omega_M^{I2} - 20\omega_M^3\omega_M^{II}) \quad . \quad . \quad (8)$$

If ω_M has the form given by eqn. (2) and $\check{\omega}_m \ll \hat{\omega}_m$, ω_M^4 is contained in ω_M^6 with a multiplying factor $15\omega_\Delta^2$ (Reference 3 and Reference 15, Section 8.2). Thus, the fourth-order distortion terms become

$$\frac{b_4'}{\omega_c^4}\omega_M^4 + 15\omega_\Delta^2\frac{b_6'}{\omega_c^6}\omega_M^4 - 45\frac{b_6'}{\omega_c^6}\omega_M^2\omega_M^{I2} - 20\frac{b_6'}{\omega_c^6}\omega_M^3\omega_M^{II}$$
$$= \frac{b_6'}{\omega_c^6}[A\omega_M^4 - 45\omega_M^2\omega_M^{I2} - 20\omega_M^3\omega_M^{II}] \quad . \quad (9)$$

In view of eqn. (2), this is

$$\frac{b_6'}{\omega_c^6}\Big[A\sum_{\omega_q=1}^{\hat{\omega}_m} a\cos(\omega_q t+\phi_{\omega q})\sum_{\omega_r=1}^{\hat{\omega}_m} a\cos(\omega_r t+\phi_{\omega r})$$
$$\sum_{\omega_s=1}^{\hat{\omega}_m} a\cos(\omega_s t+\phi_{\omega s})\sum_{\omega_t=1}^{\hat{\omega}_m} a\cos(\omega_t t+\phi_{\omega t})$$
$$- 45\sum_{\omega_q=1}^{\hat{\omega}_m} a\cos(\omega_q t+\phi_{\omega q})\sum_{\omega_r=1}^{\hat{\omega}_m} a\cos(\omega_r t+\phi_{\omega r})$$
$$\sum_{\omega_s=1}^{\hat{\omega}_m} a\omega_s\sin(\omega_s t+\phi_{\omega s})\sum_{\omega_t=1}^{\hat{\omega}_m} a\omega_t\sin(\omega_t t+\phi_{\omega t})$$
$$+ 20\sum_{\omega_q=1}^{\hat{\omega}_m} a\cos(\omega_q t+\phi_{\omega q})\sum_{\omega_r=1}^{\hat{\omega}_m} a\cos(\omega_r t+\phi_{\omega r})$$
$$\sum_{\omega_q=1}^{\hat{\omega}_m} a\cos(\omega_s t+\phi_{\omega s})\sum_{\omega_t=1}^{\hat{\omega}_m} a\omega_t^2\cos(\omega_t t+\phi_{\omega t}) \quad . \quad (10)$$

$\check{\omega}_m$ has been taken here as much less than $\hat{\omega}_m$, and put equal to unity. Expanding expression (10), the fourth-order distortion becomes

$$\frac{a^4}{8}\frac{b_6'}{\omega_c^6}\sum_{\omega_q=1}^{\hat{\omega}_m}\sum_{\omega_r=1}^{\hat{\omega}_m}\sum_{\omega_s=1}^{\hat{\omega}_m}\sum_{\omega_t=1}^{\hat{\omega}_m}\{[A + 45(\pm\omega_s)(\pm\omega_t)$$
$$+ 20\omega_r^2]\cos[(\omega_q \pm \omega_r \pm \omega_s \pm \omega_t)t + \phi_{\omega q} \pm \phi_{\omega r} \pm \phi_{\omega s} \pm \phi_{\omega t}]\}$$
$$. \quad . \quad . \quad . \quad (11)$$

where the signs are taken consistently in all possible arrangements.

To evaluate the spectral distribution of expression (9) we have to put the combination frequency $\omega_q \pm \omega_r \pm \omega_s \pm \omega_t = \omega_m$, where ω_m is a general base-band frequency. Many of the tones in expression (11) add in coherent groups. To take account of this the summations are rewritten with $\omega_q > \omega_r > \omega_s > \omega_t$ to ensure that the tones then appearing are mutually incoherent. If in expression (11) the full ranges of the variable frequencies were replaced by ranges restricted in this way, with no other modification, most of the tones would be excluded. All excluded tones, however, would be coherent with one or other of the retained tones. Thus, it is possible to modify the expression for the tone amplitudes to take account of the coherent groups. The procedure is considered in detail in Section 8.2 of Reference 15.

It is found that, after grouping coherent terms, expression (becomes

$$\frac{a^4}{8}\frac{b_6'}{\omega_c^6}\sum_{\omega_q=4}^{\hat{\omega}_m}\sum_{\omega_r=3}^{\omega_q}\sum_{\omega_s=2}^{\omega_r}\sum_{\omega_t=1}^{\omega_s}(\{24A + 120(\omega_q^2+\omega_r^2+\omega_s^2+\omega_t^2)$$
$$+ 180[\omega_q(\pm\omega_r) + \omega_q(\pm\omega_s) + \omega_q(\pm\omega_t) + (\pm\omega_r)(\pm\omega_s)$$
$$+ (\pm\omega_r)(\pm\omega_t) + (\pm\omega_s)(\pm\omega_t)]\}$$
$$\cos[(\omega_q \pm \omega_r \pm \omega_s \pm \omega_t)t + \phi_{\omega q} \pm \phi_{\omega r} \pm \phi_{\omega s} \pm \phi_{\omega t}])$$
$$. \quad . \quad . \quad . \quad (1$$

where again the signs are taken consistently in all possi arrangements.

Since all the tones in expression (12) are now incoherent, t total contribution at frequency ω_m can be found by taking t root of the sum of the squares of the amplitudes of all ton such that $\omega_q \pm \omega_r \pm \omega_s \pm \omega_t = \omega_m$ or $-\omega_m$. This conditi implies limitations on the upper limits of the summations, whi will take quite complicated forms dependent on the signs ω_r, ω_s and ω_t (cf. Reference 15, Section 8.2, where the simpl third-order distortion case is examined in detail). To avo such considerations we allow ω_q, ω_r, ω_s and ω_t to cover the full range while introducing a discontinuous factor which w count the contribution of a tone if its frequency is ω_m, or $-\omega$ and reject it otherwise. The factor required is evidently t Dirac δ function, which might, for example, be introduced its Fourier integral form,[2] i.e.

$$\delta(x-p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varepsilon^{j(x-p)z}dz$$

or as an integral of Bessel functions [obtained by differentiatir the Weber step function (p. 406 of Reference 20)]. For t present purpose, the most economical course (see Section 6. appears to be to take the discontinuous factor in the form of contour integral previously used by Bennett (p. 609 Reference 1). It can be shown, following the same line reasoning as in Reference 1, that the total contribution frequency ω_m which can be extracted from expression (12) is

$$\frac{a^4}{8}\frac{b_6'}{\omega_c^6}F(\omega_m, \hat{\omega}_m)$$

where

$$[F(\omega_m, \hat{\omega}_m)]^2 = \frac{1}{2\pi}\int_c \exp(-j\omega_m z)dz\int_0^{\hat{\omega}_m}d\omega_q\int_0^{\omega_q}d\omega_r\int_0^{\omega_r}d\omega_s$$
$$\int_0^{\omega_s}\{24A + 120(\omega_q^2+\omega_r^2+\omega_s^2+\omega_t^2) + 180[\omega_q(\pm\omega_r)$$
$$+ \omega_q(\pm\omega_s) + \omega_q(\pm\omega_t) + (\pm\omega_r)(\pm\omega_s) + (\pm\omega_r)(\pm\omega_t)$$
$$+ (\pm\omega_s)(\pm\omega_t)]\}^2\exp[j(\omega_q \pm \omega_r \pm \omega_s \pm \omega_t)z]d\omega$$
$$+ \frac{1}{2\pi}\int_c \exp(-j\omega_m z)dz\int_0^{\hat{\omega}_m}d\omega_q\int_0^{\omega_q}d\omega_r\int_0^{\omega_r}d\omega_s$$
$$\int_0^{\omega_s}\{24A + 120(\omega_q^2+\omega_r^2+\omega_s^2+\omega_t^2) + 180(-\omega_q\omega_r$$
$$- \omega_q\omega_s + \omega_q\omega_t + \omega_r\omega_s - \omega_r\omega_t - \omega_s\omega_t)\}^2$$
$$\exp[j(-\omega_q + \omega_r + \omega_s - \omega_t)z]d\omega_t + \frac{1}{2\pi}\int_c \exp(-j\omega_m z)$$
$$\int_0^{\hat{\omega}_m}d\omega_q\int_0^{\omega_q}d\omega_r\int_0^{\omega_r}d\omega_s\int_0^{\omega_s}\{24A + 120(\omega_q^2+\omega_r^2+\omega_s^2+\omega_t^2)$$
$$+ 180(-\omega_q\omega_r - \omega_q\omega_s - \omega_q\omega_t + \omega_r\omega_s + \omega_r\omega_t + \omega_s\omega_t)$$
$$\exp[j(-\omega_q + \omega_r + \omega_s + \omega_t)z]d\omega_t \quad . \quad . \quad . \quad . \quad . \quad (1$$

Table 1.—Distortion/Signal (Voltage Ratio) in the Top Channel (White-Noise Frequency Modulation)

Distortion order	Amplifier characteristics	Discriminator characteristics
Second	*Amplitude characteristic* $\frac{3}{2}b_3\frac{\omega_\Delta\hat{\omega}_m^2}{\omega_c^3}$	$\frac{1}{\sqrt{2}}\frac{\omega_c}{B_1}\omega_\Delta\Big[\Big(\frac{b_6'}{\omega_c^6}\Big)^2\Big(\frac{13\,387}{1\,260}\hat{\omega}_m^8+\frac{6\,263}{21}\omega_\Delta^2\hat{\omega}_m^6+2\,526\omega_\Delta^4\hat{\omega}_m^4+5\,850\omega_\Delta^6\hat{\omega}_m^2+4\,050\omega_\Delta^8\Big)$ $+\Big(\frac{b_4'}{\omega_c^4}\Big)^2\Big(\frac{101}{15}\hat{\omega}_m^4+44\omega_\Delta^2\hat{\omega}_m^2+72\omega_\Delta^4\Big)+2\Big(\frac{b_2'}{\omega_c^2}\Big)^2$ $+\Big(\frac{b_2'}{\omega_c^2}\Big)\Big(\frac{b_6'}{\omega_c^6}\Big)\Big(\frac{137}{15}\hat{\omega}_m^4+130\omega_\Delta^2\hat{\omega}_m^2+180\omega_\Delta^4\Big)+\Big(\frac{b_2'}{\omega_c^2}\Big)\Big(\frac{b_4'}{\omega_c^4}\Big)\Big(\frac{22}{3}\hat{\omega}_m^2+24\omega_\Delta^2\Big)$ $+\Big(\frac{b_4'}{\omega_c^4}\Big)\Big(\frac{b_6'}{\omega_c^6}\Big)\Big(\frac{1\,768}{105}\hat{\omega}_m^6+\frac{4\,402}{15}\omega_\Delta^2\hat{\omega}_m^4+1\,110\omega_\Delta^4\hat{\omega}_m^2+1\,080\omega_\Delta^6\Big)\Big]^{1/2}$
Third	$2(\sqrt{3})b_4\frac{\omega_\Delta^2\hat{\omega}_m^2}{\omega_c^4}$	$\frac{1}{2}\frac{b_5'}{B_1}\frac{\hat{\omega}_m^2\omega_\Delta^2}{\omega_c^4}\Big[\frac{12}{\hat{\omega}_m^4}\Big(\frac{b_3'}{b_5'}\omega_c^2+10\omega_\Delta^2\Big)^2+\frac{5}{\hat{\omega}_m^2}\Big(\frac{b_3'}{b_5'}\omega_c^2+10\omega_\Delta^2\Big)+\frac{470}{9}\Big]^{1/2}$
Fourth		$\frac{\sqrt{2}}{4}\frac{b_6'}{B_1}\frac{\hat{\omega}_m^2\omega_\Delta^3}{\omega_c^5}\Big[\frac{92}{\hat{\omega}_m^4}\Big(\frac{b_4'}{b_6'}\omega_c^2+15\omega_\Delta^2\Big)^2+\frac{986}{\hat{\omega}_m^2}\Big(\frac{b_4'}{b_6'}\omega_c^2+15\omega_\Delta^2\Big)-\frac{230\,495}{28}\Big]^{1/2}$
Fifth		$\sqrt{(55)}\frac{b_5'}{B_1}\frac{\omega_\Delta^4}{\omega_c^4}$
Sixth		$\frac{1}{4}\sqrt{(5\,046)}\frac{b_6'}{B_1}\frac{\omega_\Delta^5}{\omega_c^5}$
Second	*Phase characteristic* $\frac{1}{\sqrt{2}}\hat{\omega}_m\omega_\Delta\Big[\Big(\frac{a_6}{\omega_c^6}\Big)^2\Big(\frac{13\,387}{1\,260}\hat{\omega}_m^8+\frac{6\,263}{21}\omega_\Delta^2\hat{\omega}_m^6+2\,526\omega_\Delta^4\hat{\omega}_m^4+5\,850\omega_\Delta^6\hat{\omega}_m^2+4\,050\omega_\Delta^8\Big)$ $+\Big(\frac{a_4}{\omega_c^4}\Big)^2\Big(\frac{101}{15}\hat{\omega}_m^4+44\omega_\Delta^2\hat{\omega}_m^2+72\omega_\Delta^4\Big)+2\Big(\frac{a_2}{\omega_c^2}\Big)^2$ $+\Big(\frac{a_2}{\omega_c^2}\Big)\Big(\frac{a_6}{\omega_c^6}\Big)\Big(\frac{137}{15}\hat{\omega}_m^4+130\omega_\Delta^2\hat{\omega}_m^2+180\omega_\Delta^4\Big)+\Big(\frac{a_2}{\omega_c^2}\Big)\Big(\frac{a_4}{\omega_c^4}\Big)\Big(\frac{22}{3}\hat{\omega}_m^2+24\omega_\Delta^2\Big)$ $+\Big(\frac{a_4}{\omega_c^4}\Big)\Big(\frac{a_6}{\omega_c^6}\Big)\Big(\frac{1\,768}{105}\hat{\omega}_m^6+\frac{4\,402}{15}\omega_\Delta^2\hat{\omega}_m^4+1\,110\omega_\Delta^4\hat{\omega}_m^2+1\,080\omega_\Delta^6\Big)\Big]^{1/2}$	$\frac{3}{2}\frac{c_3'}{B_1}\frac{\omega_\Delta\hat{\omega}_m}{\omega_c^2}$
Third	$\frac{1}{2}a_5\frac{\hat{\omega}_m^3\omega_\Delta^2}{\omega_c^5}\Big[\frac{12}{\hat{\omega}_m^4}\Big(\frac{a_3}{a_5}\omega_c^2+10\omega_\Delta^2\Big)^2+\frac{5}{\hat{\omega}_m^2}\Big(\frac{a_3}{a_5}\omega_c^2+10\omega_\Delta^2\Big)+\frac{470}{9}\Big]^{1/2}$	$2\sqrt{(3)}\frac{B_4'}{B_1}\frac{\omega_\Delta^2\hat{\omega}_m}{\omega_c^3}$
Fourth	$\frac{\sqrt{2}}{4}a_6\frac{\hat{\omega}_m^3\omega_\Delta^3}{\omega_c^6}\Big[\frac{92}{\hat{\omega}_m^4}\Big(\frac{a_4}{a_6}\omega_c^2+15\omega_\Delta^2\Big)^2+\frac{986}{\hat{\omega}_m^2}\Big(\frac{a_4}{a_6}\omega_c^2+15\omega_\Delta^2\Big)-\frac{230\,495}{28}\Big]^{1/2}$	
Fifth	$\sqrt{(55)}a_5\frac{\omega_\Delta^4\hat{\omega}_m}{\omega_c^5}$	
Sixth	$\frac{1}{4}\sqrt{(5\,046)}a_6\frac{\omega_\Delta^5\hat{\omega}_m}{\omega_c^6}$	

In this expression the summations in expression (12) have been changed to integrals, and as before, the signs in the first set of multiple integrals must be taken consistently in all possible arrangements. The contour integration is taken along the real axis from $-\infty$ to ∞, with a downward indentation at the origin. In order to evaluate expression (13), the following reduction formulae are required (Reference 23, p. 127, and Reference 1, p. 609):

$$\int x^n \varepsilon^{ax} dx = \varepsilon^{ax}\left[\frac{x^n}{a} - \frac{nx^{n-1}}{a^2} + \frac{n(n-1)x^{n-2}}{a^3} + \ldots + (-1)^{n-1}\frac{(n)!}{a^n}x + (-1)^n\frac{(n)!}{a^{n+1}}\right] \quad . \quad (14)$$

and

$$\left.\begin{aligned}\int_c \frac{\varepsilon^{jxz}}{z^m}dz &= \frac{2\pi j^m x^{m-1}}{(m-1)!} \quad \text{for } x \geqslant 0\\ &= 0 \quad \text{for } x < 0\end{aligned}\right\} \quad . \quad . \quad (15)$$

The expressions in the curly brackets in each of the ten integrals of expression (13) must first be squared and the corresponding terms collected. In the present example this was done by hand, although in a higher-order case it would be worth considering the use of a computer, even at this stage. The successive integrations, up to the final contour integration, will now generate terms of the form

$$\alpha j^\beta A^\gamma \omega_q^\xi \omega_r^\xi \omega_s^\eta \omega_t^\theta \exp\left[j(\lambda\omega_q + \mu\omega_r + \nu\omega_s + \xi\omega_t)\right] \quad . \quad (16)$$

where the eleven Greek letters are all numbers determined by repeated application of eqn. (14).

It is apparent that the evaluation of eqn. (13) in algebraic form has been reduced to the mechanical repetition of a set of numerical operations. The number of terms to be processed would present a formidable task to a human operator (after the initial squaring, we start with 46 terms in each integral, increasing to about 130 with the first integration alone), but can be handled with no special difficulty with the help of a digital computer. Some notes on the programming procedure are given in Section 6.3. The final result emerges as

$$\begin{aligned}&\frac{\text{Distortion}}{\text{Signal}} \text{ (voltage ratio)}\\ &= \frac{\sqrt{2}}{4}\frac{b_6'}{B_1}\frac{\hat{\omega}_m^2\omega_\Delta^3}{\omega_c^5}\left[4\frac{A^2}{\hat{\omega}_m^4}\left(32 - 12\frac{\omega_m^2}{\hat{\omega}_m^2} + 3\frac{\omega_m^3}{\hat{\omega}_m^3}\right)\right.\\ &+ \frac{A}{\hat{\omega}_m^2}\left(384 + 800\frac{\omega_m^2}{\hat{\omega}_m^2} + 120\frac{\omega_m^3}{\hat{\omega}_m^3} - 420\frac{\omega_m^4}{\hat{\omega}_m^4} + 102\frac{\omega_m^5}{\hat{\omega}_m^5}\right)\\ &+ \frac{1}{28}\left(-6\,444\,385 + 15\,793\,575\frac{\omega_m}{\hat{\omega}_m} - 16\,962\,645\frac{\omega_m^2}{\omega_m^2}\right.\\ &+ 10\,818\,675\frac{\omega_m^3}{\omega_m^3} - 4\,405\,275\frac{\omega_m^4}{\omega_m^4} + 1\,109\,325\frac{\omega_m^5}{\hat{\omega}_m^5}\\ &\left.\left. - 147\,385\frac{\omega_m^6}{\hat{\omega}_m^6} + 7\,620\frac{\omega_m^7}{\hat{\omega}_m^7}\right)\right]^{1/2} = \frac{\sqrt{2}}{4}\frac{b_6'}{B_1}\frac{\hat{\omega}_m^2\omega_\Delta^3}{\omega_c^5}\\ &\left(92\frac{A^2}{\hat{\omega}_m^4} + 986\frac{A}{\hat{\omega}_m^2} - \frac{230\,495}{28}\right)^{1/2} \text{ in top channel} \quad . \quad . \quad (17)\end{aligned}$$

(4) COLLECTED EXPRESSIONS FOR DISTORTION IN TOP CHANNEL

In Table 1 are collected the formulae for the various orders of distortion in the top channel (where distortion is generally worst), so far as they are at present known. Fourth-order distortion due to the discriminator amplitude characteristic was evaluated in the previous Section; second- and third-order distortion from the same source was evaluated by hand, using the method given in Reference 15. The corresponding orders of distortion due to the amplifier phase characteristic can be deduced by noticing that the expression in the first square bracket of expression (7) is (apart from nomenclature changes and changes of sign) the time derivative of the expression in the second square bracket. This Table extends Table 1 of Reference 15—in which there is a misprint in the fourth entry from the top, where 0·47 should be 0·047.

(5) REFERENCES

(1) BENNETT, W. R.: 'Cross-Modulation Requirements of Multichannel Amplifiers Below Overload', *Bell System Technical Journal*, 1940, **19**, p. 587.

(2) LIGHTHILL, M. J.: 'Introduction to Fourier Analysis and Generalised Functions' (Cambridge University Press, 1958).

(3) LEWIN, L.: 'Interference in Multi-Channel Circuits: Dependence on Harmonic Generation', *Wireless Engineer*, 1950, **27**, p. 294.

(4) BENNETT, W. R., CURTIS, H. E., and RICE, S. O.: 'Intelligible channel Interference in F.M. and P.M. Systems Under Noise Loading Conditions', *Bell System Technical Journal*, 1955, **34**, p. 601.

(5) BORODICH, S. V.: 'On the Nonlinear Distortions caused by Variations of the Antenna Feeder in Multichannel Frequency Modulation Systems', *Radiotekhnika*, 1955, **10**, p. 3.

(6) MEDHURST, R. G.: 'Echo Distortion in Frequency Modulation, Frequency-Division-Multiplex Trunk Radio Systems', *Electronic and Radio Engineer*, 1959, **36**, p. 25.

(7) MEDHURST, R. G., and HODGKINSON, M.: 'Intermodulation Distortion due to Fading in Frequency-Modulation Frequency-Division Multiplex Trunk Radio Systems', *Proceedings I.E.E.*, Monograph No. 240 R, May, 195 (**104** C, p. 475).

(8) MEDHURST, R. G., HICKS, E. M., and GROSSETT, W. 'Distortion in Frequency-Division-Multiplex F.M. Systems due to an Interfering Carrier', *ibid.*, Pap No. 2565 R, May, 1958 (**105** B, p. 282).

(9) MEDHURST, R. G.: 'Harmonic Distortion of Frequency Modulated Waves by Linear Networks', *ibid.*, Pap No. 1650 R, May, 1954 (**101**, Part III, p. 171).

(10) BROWN, R. F.: 'Frequency-Modulation Distortion in Line Networks with Special Application to Minimum-Phas Type Networks', *ibid.*, Paper No. 2196 R, January, 19 (**104** B, p. 52).

(11) MEDHURST, R. G.: 'Fundamental and Harmonic Distortio of Waves Frequency Modulated with a Single Ton *ibid.*, Paper No. 3182 E, March, 1960 (**107** B, p. 155).

(12) MEDHURST, R. G.: 'R.F. Bandwidth of Frequency-Divisio Multiplex Systems using Frequency Modulation', *Pr ceedings of the Institute of Radio Engineers*, 1956, 4 p. 189.

(13) RICE, S. O.: 'Distortion in a Noise Modulated F.M. Sign by Nonlinear Attenuation and Phase Shift', *Bell Syste Technical Journal*, 1957, **36**, p. 879.

(14) MEDHURST, R. G., and SMALL, G. F.: 'Distortion Frequency-Modulation Systems due to Small Sin soidal Variations of Transmission Characteristics', *Pr ceedings of the Institute of Radio Engineers*, 1956, 4 p. 1608.

(15) MEDHURST, R. G., and HYAMSON, H. D.: 'Discriminat Distortion in Frequency-Modulation-Systems', *Proce*

ings I.E.E., Monograph No. 226 R, February, 1957 (**104** C, p. 357).

6) HOLBROOK, B. D., and DIXON, J. T.: 'Load Rating Theory for Multi-Channel Amplifiers', *Bell System Technical Journal*, 1939, **18**, p. 624.

7) BENNETT, W. R.: 'Methods of Solving Noise Problems', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 609.

8) MIDDLETON, D.: 'The Distribution of Energy in Randomly Modulated Waves', *Philosophical Magazine*, 1951, **42**, p. 689.

9) STEWART, J. L.: 'The Power Spectrum of a Carrier Frequency Modulated by Gaussian Noise', *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 1539.

0) WATSON, G. N.: 'A Treatise on the Theory of Bessel Functions' (Cambridge University Press, 1952), Second edition.

1) CHERRY, E. C., and RIVLIN, R. S.: 'Non-Linear Distortion with Particular Reference to the Theory of Frequency Modulated Waves—Part II', *Philosophical Magazine*, 1942, **33**, p. 272.

2) MEDHURST, R. G.: 'On the Response of Linear Systems to Signals Modulated in Amplitude and Frequency' (note on a paper by F. ZWEIG, P. M. SCHULTHEISS and C. A. WOGRIN), *Transactions of the Institute of Radio Engineers*, September, 1956, **CT-3**, p. 202.

3) DWIGHT, H. B.: 'Tables of Integrals and Other Mathematical Data' (Macmillan, New York, 1947), Second Edition.

4) LAWSON, J. L., and UHLENBECK, G. E.: 'Threshold Signals' (McGraw-Hill, 1950), Chapter 13.

(6) APPENDICES

(6.1) Derivation of Expression (7)

An expression from which the term of expression (7) in the first of square brackets can be derived is given in Reference 9. It s not then proved for a general modulation, but inferred from e form of the phase-modulation distortion when the wanted odulation is a single tone. The term of expression (7) in the cond set of square brackets was given in Reference 15, the proof volving the general expression for the spectrum of an f.m. ve with multi-tone modulation given by Cherry and Rivlin.[21] this Section an outline is given of a direct derivation of both ms.

Consider first distortion due to an amplifier characteristic, the mittance corresponding to the characteristics shown in eqns. (3) d (4) is

$$+\sum_{n=0}^{\infty}\frac{b_n}{\omega_c^n}\omega^n\Big)\exp j\Big(\sum_{n=0}^{\infty}\frac{a_n}{\omega_c^n}\omega^n\Big)$$

$$\simeq\Big(1+\sum_{n=0}^{\infty}\frac{b_n}{\omega_c^n}\omega^n+j\sum_{n=2}^{\infty}\frac{a_n}{\omega_c^n}\omega^n\Big)\exp j\Big(a_0+\frac{a_1}{\omega_c}\omega\Big)$$

hen evaluating distortion we neglect the constant and linear ms in the exponential, since these merely cause a phase shift the modulated wave and a delay of the modulation, respectively. en the output of the network can be written[22] in the form

$$\varepsilon^{j(\omega_c t+\mu_t)}\Big\{1+b_0+\frac{b_1}{\omega_c}\omega_M+\sum_{n=2}^{\infty}\Big(\frac{b_n}{\omega_c^n}+j\frac{a_n}{\omega_c^n}\Big)\Big[\frac{1}{j^n}\Big(X\frac{d^nX}{dt^n}+Y\frac{d^nY}{dt^n}\Big)+\frac{1}{j^{n+1}}\Big(Y\frac{d^nX}{dt^n}-X\frac{d^nY}{dt^n}\Big)\Big]\Big\}\quad(18)$$

where $X=\cos\mu_t$ and $Y=\sin\mu_t$.

To a first-order approximation the phase-modulation distortion is given by the imaginary term in the curly brackets. This is

$$\frac{b_2}{\omega_c^2}(YX^{II}-XY^{II})+\frac{b_3}{\omega_c^3}(XX^{III}+YY^{III})-\frac{b_4}{\omega_c^4}(YX^{IV}-XY^{IV})-\ldots-\frac{a_2}{\omega_c^2}(XX^{II}+YY^{II})+\frac{a_3}{\omega_c^3}(YX^{III}-XY^{III})+\frac{a_4}{\omega_c^4}(XX^{IV}+YY^{IV})-\ldots\quad(19)$$

where the superscripts denote differentiation with respect to time. Performing the indicated differentiations and differentiating the whole expression with respect to time to recover the frequency modulation yields the term in the first square bracket of expression (7).

The distortion due to a discriminator characteristic follows in a similar way. The admittance is now

$$\Big(B_0+\frac{B_1}{\omega_c}\omega+\sum_{n=2}^{\infty}\frac{b_n'}{\omega_c^n}\omega^n\Big)\exp j\Big(\sum_{n=2}^{\infty}\frac{a_n'}{\omega_c^n}\omega^n\Big)$$

$$\simeq B_0+\frac{B_1}{\omega_c}\omega+\sum_{n=2}^{\infty}\frac{b_n'}{\omega_c^n}\omega^n+j\Big(B_0+\frac{B_1}{\omega_c}\omega\Big)\sum_{n=2}^{\infty}\frac{a_n'}{\omega_c^n}\omega_n$$

$$=B_0+\frac{B_1}{\omega_c}\omega+\sum_{n=2}^{\infty}\frac{b_n'}{\omega_c^n}\omega^n+j\sum_{n=3}^{\infty}\frac{c_n'}{\omega_c^n}\omega^n$$

Applying the frequency-modulated wave to this admittance, the output wave can be written in a form similar to expression (18). To a first-order approximation the amplitude modulation will be given by the real term of the factor multiplying $\exp[j(\omega_c t+\mu_t)]$, in the expression obtained in this way. This turns out to be

$$B_0+\frac{B_1}{\omega_c}\omega_M-\frac{b_2}{\omega_c^2}(XX^{II}+YY^{II})+\frac{b_3}{\omega_c^3}(YX^{III}-XY^{III})+\frac{b_4}{\omega_c^4}(XX^{IV}+YY^{IV})-\ldots-\frac{c_3}{\omega_c^3}(XX^{III}+YY^{III})+\frac{c_4}{\omega_c^4}(YX^{IV}-XY^{IV})+\frac{c_5}{\omega_c^5}(XX^{V}+YY^{V})-\ldots\quad(20)$$

from which, after carrying out the differentiations, there emerges the term in the second square bracket of expression (7).

(6.2) Alternative Form of Expression (13) Using Generalized Function Theory

Considering, for example, the tones of expression (12) whose frequencies are of the form $\omega_q+\omega_r+\omega_s+\omega_t$, we require to extract tones of frequency ω_m from

$$\sum_{\omega_q=1}^{\hat{\omega}_m}\sum_{\omega_r=1}^{\omega_q}\sum_{\omega_s=1}^{\omega_r}\sum_{\omega_t=1}^{\omega_s}f(\omega_q,\omega_r,\omega_s,\omega_t)\cos[(\omega_q+\omega_r+\omega_s+\omega_t)t+\phi_{\omega_q}+\phi_{\omega_r}+\phi_{\omega_s}+\phi_{\omega_t}]$$

where $f(\omega_q, \omega_r, \omega_s, \omega_t)$ is defined in expression (12). Since all contributions will be incoherent, the total amplitude, $F(\omega_m, \hat{\omega}_m)$, of the tone of frequency ω_m is given by

$$[F(\omega_m, \hat{\omega}_m)]^2 = \int_0^{\hat{\omega}_m} d\omega_q \int_0^{\omega_q} d\omega_r \int_0^{\omega_r} d\omega_s \int_0^{\omega_s} [f(\omega_q, \omega_r, \omega_s, \omega_t)]^2 \delta(\omega_q + \omega_r + \omega_s + \omega_t - \omega_m)\, d\omega_t$$

The δ-function has the property that it is unity when $\omega_q + \omega_r + \omega_s + \omega_t = \omega_m$, and zero otherwise. Taking $\delta(\omega_q + \omega_r + \omega_s + \omega_t - \omega_m)$ in the form[2]

$$\delta(\omega_q + \omega_r + \omega_s + \omega_t - \omega_m) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp j(\omega_q + \omega_r + \omega_s + \omega_t - \omega_m) z\, dz$$

and rearranging the integral, we find that

$$[F(\omega_m, \hat{\omega}_m)]^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp - j\omega_m z \int_0^{\hat{\omega}_m} d\omega_q \int_0^{\omega_q} d\omega_r \int_0^{\omega_r} d\omega_s \int_0^{\omega_s} [f(\omega_q, \omega_r, \omega_s, \omega_t)]^2 \exp [j(\omega_q + \omega_r + \omega_s + \omega_t)z] dz \quad . \quad (21)$$

The right-hand side of eqn. (21) has a very similar appearance to that of the multiple integrals in eqn. (13), and in carrying out the evaluation the four inner integrals are, of course, dealt with in the same way. The final integration with respect to z, however, follows a different course. We find from Table 1 of Reference 2 that

$$\int_{-\infty}^{\infty} \frac{\varepsilon^{jxz}}{z^m} dz = \pi \frac{j^m x^{m-1}}{(m-1)!} \operatorname{sgn} x \quad . \quad . \quad . \quad (22)$$

where $\operatorname{sgn} x = +1$ for $x > 0$ and -1 for $x < 0$.

Eqn. (22) here replaces eqn. (15): it has the disadvanta that all terms in the final integration must be retained, where many terms drop out when eqn. (15) is used. This feature eqn. (15) has the important practical consequence (Section 6. that families of terms may be dropped in the earlier integratio Of course, the use of eqn. (22) must ultimately lead to the sa result when all terms are collected, but it appears that eqn. (1 presents the more economical course.

(6.3) Note on the Computational Procedure

The reduction to algebraic form of the ten integrals in expre sion (13) was carried out on a Hec 2M digital computer, whi uses punched-card input and output. Each integral w evaluated separately. The coding arrangements were su that each term of the form (16), specified by eleven numbe was represented by one card. Thus, prior to the first integratio it is necessary to punch up 46 cards. The programme perfor the operation indicated by expression (14), generating a new s of cards (about 130 after the first integration) of precise similar form to the input cards. Since the programme w designed to carry out each integration using a specified variab of integration, it would now be possible to return these car to the machine without further processing for the second roun of integration. If this were done, rather more than 7000 car would emerge from the fourth integration. To avoid this, t output cards at each stage were sorted in a high-speed sort into coherent groups, which were then added using a secon programme. Besides reducing the number of cards to manag able proportions, this procedure allows one to throw awa groups of terms in each stage, since, in view of the secon alternative of the final contour integration [eqn. (15)], it becom possible by inspection of the exponential index to predict th these terms will make no final contribution. After the fourt integration and a final sorting, a further programme perfor the operation indicated by eqn. (15).

1–52: 621.3.013.1

The Institution of Electrical Engineers
Monograph No. 353 M
Jan. 1960
©

AN APPROXIMATION TO THE HARMONIC RESPONSE OF SATURATING DEVICES

By R. J. KAVANAGH, B.Sc., M.A.Sc., Ph.D., Graduate.

(*The paper was first received 7th July, and in revised form 28th September*, 1959. *It was published as an* INSTITUTION MONOGRAPH *in January*, 1960.)

SUMMARY

The paper derives an analytical expression for the describing function and for the Fourier coefficients corresponding to saturating systems which are excited by sinusoidal signals. The saturation characteristics are assumed to be symmetrical and single valued. The method involves an approximation of the actual system characteristic by means of an exponential curve.

The results are given in the form of graphs of the fundamental and third-harmonic components of the system output and a graph of the system describing function. The latter will be of use in the analysis of non-linear feedback systems. A comparison is shown between the present method and the technique of approximating saturation characteristics by means of straight lines. Formulae are given by means of which the higher-order harmonic components of the system output may be computed.

Secondary results of this paper are Tables of the modified Struve functions $L_1(z)$ and $L_2(z)$ which are not available elsewhere in the literature.

LIST OF SYMBOLS

A = Output of a system when completely saturated.
$a_m(z)$ = $(m + 1)$th term of the series for $L_\nu(z)$.
a_n = Fourier coefficient for the nth harmonic of the system output.
B = Amplitude of the system input sine wave.
c = Constant such that Ac is the small-signal gain of the system.
$I_\nu(z)$ = Modified Bessel function of the first kind of order ν for an argument z.
k, m, n = Summation variables.
$L_\nu(z)$ = Modified Struve function of order ν for an argument z.
N = System describing function.
P = Integral involved in the derivation of Fourier coefficients.
x, y = The system input and output variables.
$\Gamma(z)$ = Gamma function for an argument z.
θ = System input angle.

(1) INTRODUCTION

A commonly occurring phenomenon in all types of physical system is that of saturation. In the field of electrical engineering, for example, saturation of electronic amplifiers, magnetic saturation and dielectric saturation are commonly encountered. The analysis of systems containing such non-linearities is necessarily difficult since the principle of superposition is not valid.

In electrical systems, devices having saturation characteristics are often subjected to sinusoidal excitation. As a result, the output of the device will contain harmonic components as well as a component having the same frequency as the input excitation. A knowledge of the amplitudes of these various components of the output may be of considerable interest as, for example, in a.c. machines. In the field of feedback control the analysis of systems which contain saturating devices is commonly accomplished by the use of a describing function,[1] which, for a saturating device, and also for any other non-linearity, is defined as the ratio of the amplitude of the fundamental component of the output, expressed as a complex number, to that of the input. It is assumed that the input to the non-linearity is purely sinusoidal.

Since practical saturation characteristics are not known in exact analytical form, the most accurate method of determining the amplitudes of the various components of the output in response to a sinusoidal input is to perform a Fourier analysis by a numerical method. This method has the disadvantage of being quite tedious even if a relatively small number of ordinates is used. An alternative method which may be employed is to approximate the actual saturation characteristic by a curve which may be described analytically. The Fourier analysis of this approximate curve may now be performed exactly. The resultant Fourier coefficients are then used as approximations to those corresponding to the original saturation characteristic.

The paper is concerned with an approximation method of this type which is able to give a close estimate of the fundamental component of the output with a minimum of labour. The method is thus well suited to the describing-function analysis of feedback control systems containing saturating devices. Calculation of the third-harmonic component of the output is obtained equally easily by use of a graph given in the paper. Higher-order harmonics may be computed using formulae which are also given. It is to be noted that the saturation characteristics discussed here are symmetrical and single valued.

(2) APPROXIMATIONS TO SATURATION CHARACTERISTICS

There are many analytically describable curves which could be used to approximate an actual characteristic. While the paper is concerned with one particular curve, it would seem worth while to consider several other possible curves with a view to comparing their utility.

Four possible curves will be considered, namely

(*a*) 3-segment straight line.

$$\left.\begin{aligned} y &= Acx \qquad -\frac{1}{c} \leqslant x \leqslant +\frac{1}{c} \\ y &= A \qquad x > \frac{1}{c} \\ y &= -A \qquad x < -\frac{1}{c} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (1)$$

(*b*) Hyberbolic tangent.

$$y = A \tanh cx \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Correspondence on Monographs is invited for consideration with a view to publication.

Dr. Kavanagh was formerly Lecturer in Electrical Engineering, University of Toronto, and is now in the Electrical Engineering Department, Imperial College of Science and Technology, University of London.

(*c*) Froelich's equation (hyperbola).

$$\left.\begin{aligned} y &= \frac{Ax}{\frac{1}{c}+x} \quad x \geqslant 0 \\ y &= \frac{Ax}{\frac{1}{c}-x} \quad x \leqslant 0 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

(*d*) Exponential.

$$y = \pm A[1 - \exp(-c|x|)] \quad . \quad . \quad . \quad . \quad (4)$$

where $y > 0 \quad x > 0$

$y < 0 \quad x < 0$

It will be noticed that the four curves considered have been chosen such that only two parameters must be specified in order to define a particular curve. These two parameters are A, the maximum value of the system output y, and c, where Ac is the slope of each of these curves at the origin. Graphs of each of the four curves are shown in Fig. 1 for positive values of x, the independent variable, and for $A = 1$, $c = 5$.

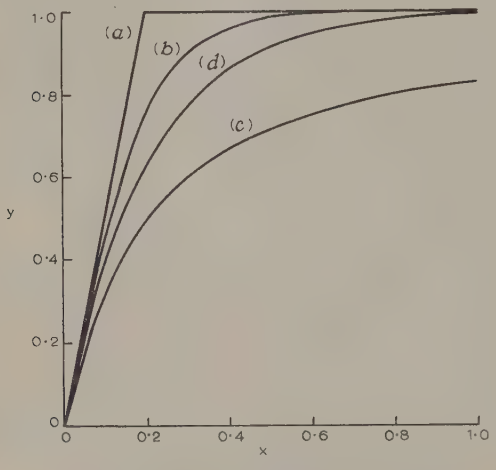


Fig. 1.—Comparison of approximating curves.
(*a*) Straight line.
(*b*) $y = \tanh 5x$.
(*c*) $y = x/0{\cdot}2 + x$.
(*d*) $y = 1 - \exp(-5x)$.

Because only two parameters are needed to define a specific curve, it is relatively easy to approximate the given saturation curve by one of the curves shown in Fig. 1. Knowledge of the maximum value of the system output and of the system gain for very small signals are sufficient to determine a specific curve, although in some cases, as discussed in Section 5, the value of Ac may be chosen not exactly equal to the slope of the characteristic at the origin. The use of approximating curves which require more than two parameters to be determined will enable a closer approximation to be made, although, of course, this will be at the expense of ease of analysis.

A comparison of the curves in Fig. 1 shows that the straight-line approximation will be most suitable for systems which have very sudden limiting of their output. This approximation for a saturation characteristic is widely used in the describing-function analysis of non-linear feedback control systems. The Fourier coefficients for the fundamental and third-harmonic components of the output of a system having the straight-line approximation and excited by a sinusoidal input a readily available in the literature.[2]

The hyperbolic-tangent approximation gives a curve wit relatively rapid saturation and hence could be a useful approx mation for many physical systems. To the author's knowled the appropriate Fourier coefficients for a system having hyberbolic-tangent characteristic are not available in the liter ture. If this is the case, it would seem worth while to deri such coefficients.

Froelich's equation is well known as an approximation fo magnetization curves. As may be seen from Fig. 1, the curv approaches saturation very slowly. A method of obtainin the Fourier coefficients corresponding to this curve by a series summation method is given in Reference 3, which also lis some additional 2-parameter approximating curves for satura tion characteristics.

The exponential approximation, which is to be discussed i the paper, is seen to lie between Froelich's equation and th hyperbolic tangent curve so far as rapidity of saturation i concerned. This approximation is suggested in Reference 3, bu no Fourier coefficients in closed form are given in the literatur to the knowledge of the author. This approximating curve i of a shape which makes it useful for the approximation of man practical saturation characteristics. An example of its use fo the approximation of a typical electronic-amplifier characteristic is given in Section 5.

(3) DERIVATION OF FOURIER COEFFICIENTS

From eqn. (4) the saturation characteristic is to be described by

$$y = \pm A[1 - \exp(-c|x|)] \quad . \quad . \quad . \quad . \quad (5)$$

where $y > 0 \quad x > 0$

$y < 0 \quad x < 0$

For a sinusoidal excitation

$$x = B\cos\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

The Fourier series for $y(\theta)$ is of the form

$$y(\theta) = a_1\cos\theta + a_3\cos 3\theta + \ldots \quad . \quad . \quad . \quad (7)$$

The Fourier coefficients may be evaluated by consideration of that portion of $y(\theta)$ which lies between $\theta = 0$ and $\theta = \pi/2$.

In this region

$$y(\theta) = A[1 - \exp(-cB\cos\theta)] \quad . \quad . \quad . \quad (8)$$

Hence the Fourier coefficients are given by

$$a_n = \frac{4A}{\pi}\int_0^{\pi/2}[1 - \exp(-cB\cos\theta)]\cos n\theta d\theta \quad . \quad . \quad (9)$$

where n is odd. It is shown in Section 8.1 that a_n may be expressed in terms of the following finite series in which $z = cB$:

$$a_n = 2A[\mathrm{I}_1(z) - \mathrm{L}_1(z)] + \frac{4A}{\sqrt{\pi}}\sum_{k=4,6..}^{n+1}(-1)^{(k+2)/2} \times \frac{(n^2-1^2)(n^2-3^2)\ldots[n^2-(k-3)^2]\Gamma\left(\frac{k+1}{2}\right)}{(k-1)!\left(\frac{z}{2}\right)^{(k-2)/2}} \times [\mathrm{I}_{k/2}(z) - \mathrm{L}_{k/2}(z)] \quad . \quad (10)$$

Expressions for the first- and third-harmonic coefficients are
en

$$a_1 = 2A[I_1(cB) - L_1(cB)] \quad . \quad . \quad . \quad . \quad (11)$$

$$a_3 = 2A[I_1(cB) - L_1(cB)] - \frac{8A}{cB}[I_2(cB) - L_2(cB)]$$

$$= a_1 - \frac{8A}{cB}[I_2(cB) - L_2(cB)] \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

he describing function for the non-linearity is then given by

$$N = \frac{a_1}{B} = \frac{2A}{B}[I_1(cB) - L_1(cB)] \quad . \quad . \quad . \quad (13)$$

(4) COMPUTATION OF FOURIER COEFFICIENTS

In order to compute the Fourier coefficients, values for the unctions $L_{k/2}$ and $I_{k/2}$ are required for a suitable range of the rgument $z = cB$. Since the coefficients a_1 and a_3 are of greatest nterest, computation of $L_1(z)$, $L_2(z)$, $I_1(z)$ and $I_2(z)$ only is equired here. As shown below, values of the higher-order unctions required for the computation of further Fourier oefficients may be readily obtained by means of recursion ormulae or by summation of a series.

The only modified Struve functions which appear to be abulated in the literature are $L_{-2}(z)$, $L_{-1}(z)$ and $L_0(z)$, as given in Reference 4. Now in general it is true that

$$L_{\nu-1}(z) - L_{\nu+1}(z) = \frac{2\nu}{z}L_\nu(z) + \left[\left(\frac{z}{2}\right)^\nu \Big/ \Gamma(\nu + \tfrac{3}{2})\Gamma(\tfrac{1}{2})\right] \quad . \quad (14)$$

If $\nu = 0$, this reduces to

$$L_1(z) = -\frac{2}{\pi} + L_{-1}(z) \quad . \quad . \quad . \quad . \quad (15)$$

and if $\nu = 1$,

$$L_2(z) = L_0(z) - \frac{2L_1(z)}{z} - \frac{2z}{3\pi} \quad . \quad . \quad . \quad (16)$$

Using eqn. (15), Table 1 of $L_1(z)$ was computed for z ranging from 0 to 5 in intervals of 0·1. $z = 5$ corresponds to the output of the non-linearity reaching to within about 0·7% of its saturated value. Because the values of $L_{-1}(z)$ used were tabulated to 8 decimals, so also were those of $L_1(z)$. The maximum error is ± 1 in the last figure.

Values of $L_2(z)$ shown in Table 2 were computed for the same range of z using eqn. (16), Table 1 and the Table of $L_0(z)$ in Reference 4. Because the operations of division and subtraction in eqn. (16) can lead to loss of several significant figures, $L_2(z)$ was rounded off to four decimal places which should be sufficient for most engineering work.

In order to compute $L_\nu(z)$ for $\nu > 2$, the recursion formula of eqn. (14) may be used or the following series[4] may be summed:

$$L_\nu(z) = \sum_{m=0}^{\infty} a_m(z) \quad . \quad . \quad . \quad . \quad . \quad (17)$$

where

$$a_m(z) = \frac{(\frac{1}{2}z)^{2m+1+\nu}}{\Gamma(m + \frac{3}{2})\Gamma(m + \nu + \frac{3}{2})} \quad . \quad . \quad . \quad (18)$$

Values of the function $I_\nu(z)$ are more readily obtainable. A useful Table for the purposes of this paper is given by Dwight,[5] where $I_\nu(z)$ is tabulated to five significant figures for integral values of ν from 0 to 11 and z from 0 to 6 in intervals of 0·1.

Calculated values for the Fourier coefficients a_1 and a_3 are given in graphical form in Fig. 2 where a_1/A and a_3/A are plotted against cB. These graphs should be sufficiently accurate to be

Table 1

VALUES OF THE MODIFIED STRUVE FUNCTION $L_1(z)$

z	$L_1(z)$	z	$L_1(z)$
0·0	0·000 000 00	2·6	2·220 227 88
0·1	0·002 123 48	2·7	2·474 943 22
0·2	0·008 510 93	2·8	2·754 310 21
0·3	0·019 213 48	2·9	3·060 674 15
0·4	0·034 316 88	3·0	3·396 612 95
0·5	0·053 942 19		
0·6	0·078 246 81	3·1	3·764 959 82
0·7	0·107 425 88	3·2	4·168 828 31
0·8	0·141 713 92	3·3	4·611 639 84
0·9	0·181 386 85	3·4	5·097 153 98
1·0	0·226 764 38	3·5	5·629 501 79
1·1	0·278 212 75	3·6	6·213 222 46
1·2	0·336 147 87	3·7	6·853 303 67
1·3	0·401 038 94	3·8	7·555 225 94
1·4	0·473 412 49	3·9	8·325 011 47
1·5	0·553 856 91	4·0	9·169 277 81
1·6	0·643 027 57	4·1	10·095 296 99
1·7	0·741 652 46	4·2	11·111 060 45
1·8	0·850 538 57	4·3	12·225 350 61
1·9	0·970 578 79	4·4	13·447 819 51
2·0	1·102 759 79	4·5	14·789 075 37
2·1	1·248 170 52	4·6	16·260 777 81
2·2	1·408 011 74	4·7	17·875 742 64
2·3	1·583 606 50	4·8	19·648 057 18
2·4	1·776 411 69	4·9	21·593 207 04
2·5	1·988 030 77	5·0	23·728 215 78

Table 2

VALUES OF THE MODIFIED STRUVE FUNCTION $L_2(z)$

z	$L_2(z)$	z	$L_2(z)$
0·0	0·000 00	2·6	1·025 5
0·1	0·000 04	2·7	1·177 1
0·2	0·000 34	2·8	1·346 7
0·3	0·001 15	2·9	1·536 2
0·4	0·002 7	3·0	1·747 7
0·5	0·005 4		
0·6	0·009 3	3·1	1·983 3
0·7	0·014 9	3·2	2·245 6
0·8	0·022 4	3·3	2·537 3
0·9	0·032 2	3·4	2·861 4
1·0	0·044 5	3·5	3·221 3
1·1	0·059 8	3·6	3·620 5
1·2	0·078 5	3·7	4·063 0
1·3	0·101 0	3·8	4·553 4
1·4	0·127 8	3·9	5·096 5
1·5	0·159 4	4·0	5·697 6
1·6	0·196 3	4·1	6·362 6
1·7	0·239 1	4·2	7·098 2
1·8	0·288 6	4·3	7·911 3
1·9	0·345 4	4·4	8·810 1
2·0	0·410 3	4·5	9·803 0
2·1	0·484 2	4·6	10·899 9
2·2	0·568 0	4·7	12·111 1
2·3	0·662 9	4·8	13·448 3
2·4	0·769 9	4·9	14·924 5
2·5	0·890 3	5·0	16·553 6

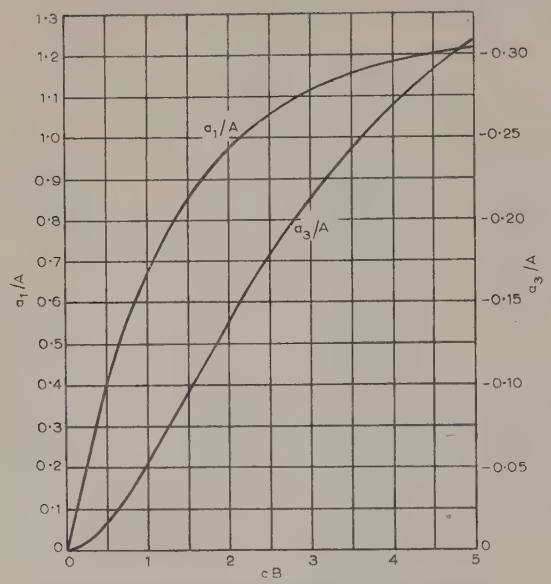


Fig. 2.—Graphs of fundamental and third harmonic.

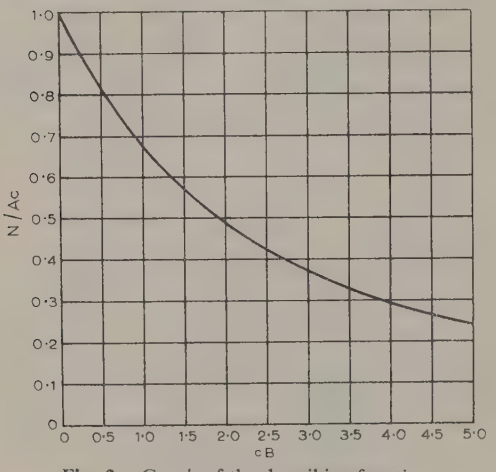


Fig. 3.—Graph of the describing function.

useful for engineering purposes. For use in the analysis of non-linear feedback control systems, a plot of the variation of N/Ac with cB is given in Fig. 3. This graph should have adequate accuracy for such a purpose.

(5) CURVE FITTING

The utility of the method described in the paper is dependent on a procedure for approximating the actual characteristic by the exponential curve. A description of the curve-fitting procedure in an actual application of the method will reveal many of the points to be considered.

The problem to be considered is that of obtaining a plot of the variation of the describing function with the input amplitude for an electronic amplifier whose characteristic curve is shown in Fig. 4. Only the curve for positive x is shown here. The corresponding curve for negative x was symmetrical with differences of such small magnitude as to be negligible. By actual measurement from the characteristic it is found that $A = 47$ and $c = 3{\cdot}7$.

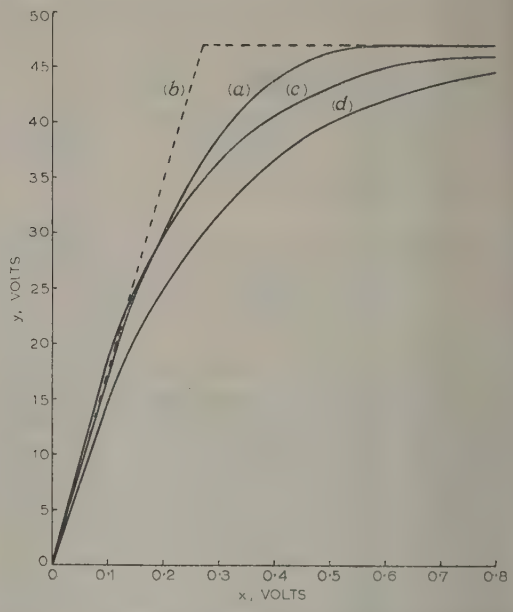


Fig. 4.—Various approximations to the amplifier characteristic.

(*a*) Amplifier characteristic.
(*b*) Straight line, $c = 3{\cdot}7$.
(*c*) Exponential, $c = 5$.
(*d*) Exponential, $c = 3{\cdot}7$.

The selection of a suitable exponential curve must now be considered. The logical choice would appear to be an exponential curve for which $A = 47$ and $c = 3{\cdot}7$. This is a characteristic such that its small-signal gain, namely 174, is the same as that of the actual amplifier. However, by reference to Fig. 4, it is seen that this exponential is not a very good fit in the sense of maximum vertical distance between it and the amplifier characteristic. Whether the exponential having the same values of A and c as the actual curve is a good fit depends entirely on the shape of the actual curve. If the curve had, by chance, been of exponential shape a perfect fit could, of course, be obtained. It will thus be seen that those characteristics which are best approximated by the methods of the paper will be those having approximately exponential shape.

Consideration will now be given to the question of whether a better fit, in the sense mentioned above, could be obtained by an exponential curve with $A = 47$ but with a value of c other than $3{\cdot}7$. Since this question is likely to arise in most applications of the method of the paper, curves of $1 - \exp(cx)$ versus x have been given in Fig. 5 for various values of c. To use these curves it is suggested that the normalized system output (i.e. actual output divided by the saturation-level output) be plotted against the normalized system input (i.e. actual input divided by the maximum input) on tracing paper to the same scale as that of Fig. 5. By superimposing this trace on Fig. 5, it is a fairly simple matter to pick a suitable value for c.

In Fig. 4 are plotted exponentials with $A = 47$ and $c = 5$ and $c = 3{\cdot}7$, and the original characteristic curve. The fit between the exponential with $c = 5$ and the characteristic curve is judged to be about the best. The question of how logically to choose the best exponential is a difficult one to answer. The criterion by which the fit is judged must surely be some function of the difference between the Fourier coefficients corresponding to the approximating curve and those corresponding to the actual curve. How this function (if a suitable one was chosen) could be related to the fit between the two curves is a problem which

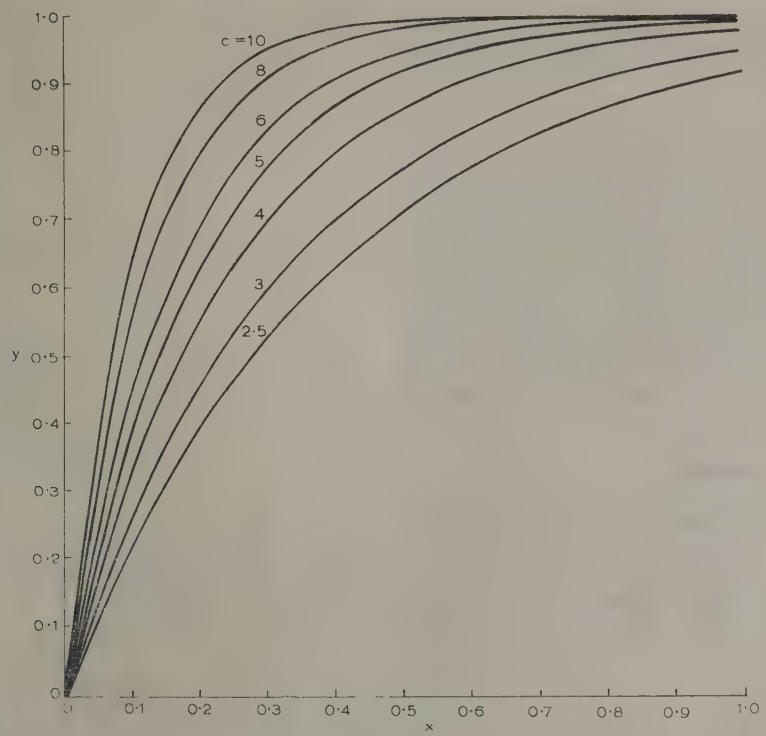


Fig. 5.—Exponential curves for various values of c.
$y = 1 - \exp(-cx)$.

he paper does not attempt to solve. In choosing the value of = 5, however, the following intuitive criteria were used:

(*a*) The vertical distance between the two curves must be kept o a minimum at all points.

(*b*) The net area between the two curves should be as small as ossible.

(*c*) The slopes of the curves near the origin should be as close s possible.

The results obtained using the exponential with $c = 5$ seem to ustify this approach and it is suggested that, in lieu of any better riteria, the above method will give satisfactory results in most ases.

The variation of the describing function N with B is given in ig. 6 for the actual curve and for various approximations to the urve. The describing function for the true curve was computed y a numerical Fourier analysis, and those for the two exponenal curves with $c = 3{\cdot}7$ and $c = 5$ were computed using Fig. 3. he difference in effort required to obtain N by these two methods s an indication of the usefulness of the approximation technique. lso shown in Fig. 6 is the describing function for the straightne approximation shown in Fig. 4, with $c = 3{\cdot}7$, which is the pproximation commonly used in non-linear control-system nalysis.

By comparison of the various curves in Fig. 6 it may be seen hat the describing function for the exponential with $c = 5$ is efinitely much more accurate than that with $c = 3{\cdot}7$ except for ow B. A comparison of the former with the straight-line escribing function again shows that the exponential approximaion gives a considerable improvement in accuracy except for B ess than about 0·15. The reason for this discrepancy at low B s the fact that the slope of the exponential for small B is greater han that of the actual curve in this region.

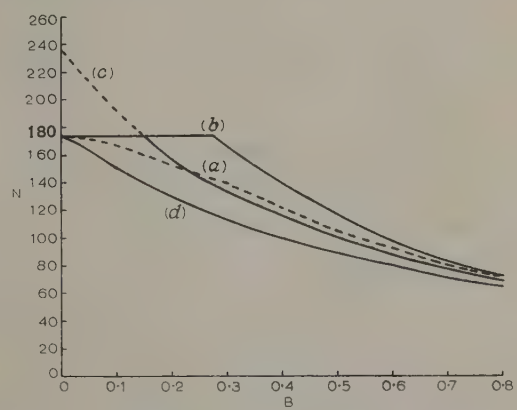


Fig. 6.—Comparison of describing functions for amplifier.
(*a*) True curve.
(*b*) Straight line, $c = 3{\cdot}7$.
(*c*) Exponential, $c = 5$.
(*d*) Exponential, $c = 3{\cdot}7$.

A great improvement in the accuracy of the describing function of the exponential for low values of B can now be obtained by recognizing that the describing function for any smoothly saturating curve cannot exceed Ac. Hence, in this case, an excellent approximation to the actual describing function is obtained by using the exponential describing function for $c = 5$ for B greater than or equal to about 0·14 and the straight-line describing function for B less than this value. The resulting 2-section describing function has a maximum deviation from the true describing function of only about 7% as compared with a deviation of about 21% for the straight-line approximation.

This procedure of constructing a 2-section describing function is always possible where the value of c chosen for the exponential using Fig. 5 is greater than that corresponding to the actual characteristic.

(6) CONCLUSIONS

The mathematical derivation of the paper results in a relatively simple method for the determination of the describing function of a physical device which is subject to saturation. The method is also useful for the computation of the Fourier components of the output of such a device when driven by a sinusoidal input.

The particular approximating curve is only one of several possibilities. The derivation of the Fourier components corresponding to alternative approximating curves would be a field for further study. In addition, an investigation of possible criteria for the curve-fitting process could yield useful results.

(7) REFERENCES

(1) KOCHENBURGER, R. J.: 'A Frequency Response Method for Analyzing and Synthesizing Contactor Servomechanisms', *Transactions of the American I.E.E.*, 1950, **69**, Part I, p. 270.

(2) TRUXAL, J. G.: 'Automatic Feedback Control System Synthesis' (McGraw-Hill, 1955), Chapter 10.

(3) HARTEL, W.: 'Abschätzung der Oberwellen bei gleichstromvormagnetisierten Drosseln', *Archiv für Elektrotechnik*, 1942, **36**, p. 556.

(4) MATHEMATICAL TABLES PROJECT: 'Table of the Struve Functions $L_\nu(x)$ and $H_\nu(x)$', *Journal of Mathematics and Physics*, 1946, **25**, p. 252.

(5) DWIGHT, H. B.: 'A Five-Figure Table of the Bessel Function $I_N(x)$', *Transactions of the American I.E.E.*, 1941, **60**, p. 135.

(6) MCLACHLAN, N. W.: 'Bessel Functions for Engineers' (Oxford, 1955), List of Formulae.

(7) WATSON, G. N.: 'A Treatise on the Theory of Bessel Functions' (Cambridge, 1922), Chapter 10.

(8) APPENDIX

Derivation of Fourier Coefficients

From eqn. (9)

$$a_n = \frac{4A}{\pi}\int_0^{\pi/2}[1 - \exp(-cB\cos\theta)]\cos n\theta d\theta \quad . \quad (19)$$

where n is odd. Separating the two terms of the integral gives

$$a_n = \frac{4A}{\pi}\int_0^{\pi/2}\cos n\theta d\theta - \frac{4A}{\pi}\int_0^{\pi/2}\exp(-cB\cos\theta)\cos n\theta d\theta \quad . \quad (20)$$

The first term of this expression when evaluated is

$$(-1)^{(n-1)/2}\frac{4A}{n\pi}$$

Consider next the integral

$$P = \int_0^{\pi/2}\exp(-cB\cos\theta)\cos n\theta d\theta \quad . \quad . \quad . \quad (21)$$

It follows that

$$P = \int_0^{\pi/2}\cosh(cB\cos\theta)\cos n\theta d\theta - \int_0^{\pi/2}\sinh(cB\cos\theta)\cos n\theta d\theta \quad . \quad . \quad . \quad . \quad (22)$$

Integrating by parts

$$P = \left[\cosh(cB\cos\theta)\frac{\sin n\theta}{n}\right]_0^{\pi/2} + cB\int_0^{\pi/2}\frac{\sin n\theta}{n}\sinh(cB\cos\theta)\sin\theta d\theta - \left[\sinh(cB\cos\theta)\frac{\sin n\theta}{n}\right]_0^{\pi/2} - cB\int_0^{\pi/2}\frac{\sin n\theta}{n}\cosh(cB\cos\theta)\sin\theta d\theta \quad . \quad . \quad (23)$$

In this expression the third term vanishes and the first term becomes

$$\frac{(-1)^{(n-1)/2}}{n}$$

To obtain the second and fourth terms in a more useful form consider the expansion for $\sin n\theta$ in terms of the powers of $\sin\theta$ for the case where n is odd:

$$\sin n\theta = n\sin\theta - \frac{n(n^2-1^2)}{3!}\sin^3\theta + \frac{n(n^2-1^2)(n^2-3^2)}{5!}\sin^5\theta - + \ldots \quad (24)$$

It follows that

$$\frac{\sin n\theta\sin\theta}{n} = \sin^2\theta - \frac{(n^2-1^2)}{3!}\sin^4\theta + \frac{(n^2-1^2)(n^2-3^2)}{5!}\sin^6\theta - + \ldots$$

$$(-1)^{(k+2)/2}\frac{(n^2-1^2)(n^2-3^2)\ldots[n^2-(k-3)^2]}{(k-1)!}\sin^k\theta\ldots \quad (25)$$

Hence P may now be written as

$$P = \frac{(-1)^{(n-1)/2}}{n} + cB\int_0^{\pi/2}\sinh(cB\cos\theta)\sin^2\theta d\theta + cB\sum_{k=4,6\ldots}^{n+1}\int_0^{\pi/2}(-1)^{(k+2)/2} \times \frac{(n^2-1^2)(n^2-3^2)\ldots[n^2-(k-3)^2]\sinh(cB\cos\theta)\sin^k\theta}{(k-1)!}d\theta - cB\int_0^{\pi/2}\cosh(cB\cos\theta)\sin^2\theta d\theta - cB\sum_{k=4,6\ldots}^{n+1}\int_0^{\pi/2}(-1)^{(k+2)/2} \times \frac{(n^2-1^2)(n^2-3^2)\ldots[n^2-(k-3)^2]\cosh(cB\cos\theta)\sin^k\theta}{(k-1)!}d\theta \quad . \quad . \quad . \quad . \quad (26)$$

Now the modified Bessel function of the first kind of order ν is given by[6]

$$I_\nu(z) = \frac{2(z/2)^\nu}{\sqrt{\pi}\Gamma(\nu+\frac{1}{2})}\int_0^{\pi/2}\cosh(z\cos\theta)\sin^{2\nu}\theta d\theta \quad . \quad (27)$$

ere $\mathscr{R}(\nu) > -\frac{1}{2}$. Also the modified Struve function of order is given by[7]

$$L_\nu(z) = \frac{2(z/2)^\nu}{\Gamma(\nu+\frac{1}{2})\Gamma(\frac{1}{2})}\int_0^{\pi/2}\sinh(z\cos\theta)\sin^{2\nu}\theta d\theta \quad . \quad (28)$$

nce, by rearrangement and comparison with the above expressns, P may be written as

$$= \frac{(-1)^{(n-1)/2}}{n} + \frac{\pi}{2}L_1(cB) - \frac{\pi}{2}I_1(cB)$$

$$+ \sum_{k=4,6\ldots}^{n+1} \frac{(-1)^{(k+2)/2}(n^2-1^2)(n^2-3^2)\ldots[(n^2-(k-3)^2]cB}{(k-1)!}$$

$$\left[\int_0^{\pi/2}\sinh(cB\cos\theta]\sin^k\theta d\theta - \int_0^{\pi/2}\cosh(cB\cos\theta)\sin^k\theta d\theta\right] \quad (29)$$

$$= \frac{(-1)^{(n-1)/2}}{n} + \frac{\pi}{2}[L_1(cB) - I_1(cB)]$$

$$+ \sum_{k=4,6\ldots}^{n+1} (-1)^{(k+2)/2}\frac{(n^2-1^2)(n^2-3^2)\ldots[n^2-(k-3)^2]cB}{(k-1)!}$$

$$\left[\frac{L_{k/2}(cB)\Gamma\left(\frac{k+1}{2}\right)\Gamma(\frac{1}{2})}{2\left(\frac{cB}{2}\right)^{k/2}} - \frac{I_{k/2}(cB)\sqrt{\pi}\Gamma\left(\frac{k+1}{2}\right)}{2\left(\frac{cB}{2}\right)^{k/2}}\right] \quad . \quad (30)$$

Substituting this expression for P in eqn. (20), and noting that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, a_n may be written as

$$a_n = 2A[I_1(cB) - L_1(cB)]$$

$$+ \frac{4A}{\sqrt{\pi}} \sum_{k=4,6\ldots}^{n+1} (-1)^{(k+2)/2}$$

$$\frac{(n^2-1^2)(n^2-3^2)\ldots[n^2-(k-3)^2]\Gamma\left(\frac{k+1}{2}\right)}{(k-1)!\left(\frac{cB}{2}\right)^{(k-2)/2}}$$

$$[I_{k/2}(cB) - L_{k/2}(cB)] \quad . \quad (31)$$

DISCUSSION ON 'EFFECTS OF ARGON CONTENT ON THE CHARACTERISTICS OF NEON-ARGON GLOW-DISCHARGE REFERENCE TUBES'*

BEFORE THE ELECTRONICS AND COMMUNICATIONS SECTION, 23RD MARCH, 1959

Mr. A. J. Young: The authors point out that the influence of rgon on glow-discharge tubes has been studied for a long time. hese effects were, of course, known in Germany in the 1930's. o far as I know, however, there has been very little published nformation before this series of papers. I think now we have omething we can use on a subject which has been studied in nore detail.

I am not surprised that the authors found great variations ith the film-coated tubes. After all, the film is very thin and hen one realizes the difference in surface conditions that one ets, even with a thermionic cathode, even greater variations ould be expected with these extremely thin, highly reactive films.

The authors conclude that the most desirable characteristics of eon tubes with molybdenum cathodes are obtained when the rgon content is about 1·0%. I do not think that is right. For xample, the plateau of the curve in Fig. 2 is obtained at about %. If the voltage is set at that point, small variations of argon ill not make very much difference. On the other hand, at 1% e are almost starting on the slippery slope.

In Fig. 3, there are no obvious jumps or steps at 1% of argon, ut after all, 0·5% is nearly as good. With that content, and the steps are caused by the discharge jumping over the edge f the cylinder, then probably they could be eliminated by djustment. As regards the temperature coefficient shown in ig. 4, we agree that 1% is the best. Against that, however, here is only a change of 200 millivolts with a 50° C change t 0·5%. In Fig. 6, initial running-voltage drift shows little ifference between 0·5 and 1·0%.

When we consider the impedance of the tube in Figs. 9 and 10, I agree that rather more argon seems to help, but against that you have to bear in mind that these are reference tubes which are not really intended as coupling elements. For that purpose, within the circuit, no doubt other conditions will apply.

In Fig. 13, minimum noise occurs at about ½%. If a reference tube is to be made using the information in the paper, perhaps 1% is a little too high. Of course, at a higher content, of the order of 3% or more, the tube would clearly not be satisfactory.

Mr. J. M. Glackin: In considering the best temperature coefficient there are two applications to be borne in mind. The first is where large changes in ambient temperature may occur. The change in maintaining voltage may be compensated for by the use of temperature-dependent circuit-elements. The temperature coefficient should therefore be as linear as possible and reproducible from tube to tube. Secondly, where very high stability is required and only small variations of temperature may occur, there should be no sudden changes in maintaining voltage due to these small changes in ambient temperature, i.e. no 'thermal jumps'.

Can the authors give more details of their methods of measurement of temperature coefficient and comment on the discontinuities in these coefficients?

Geometry has a large effect on all the tube parameters mentioned, particularly on the noise content. I have found that the level of gas noise is a function of gas pressure, anode area relative to anode-cathode spacing, and current.

There is a small but significant difference in the authors' results compared to those of Penning and others. Could this be due to a change in gas mixture caused by a preferential

* Benson, F. A., and Chalmers, P. M.: Monograph No. 321 R, December, 1958 ee 106 C, p. 82).

gettering of one of the constituent gases during the very heavy sputtering used in processing this type of tube?

Mr. J. Smith: The authors say that temperature coefficients can be explained by maintaining-voltage/pressure curves. Such a curve having a negative gradient could explain a negative temperature coefficient because, when the envelope temperature is increased, the density in the cathode dark space is increased. The effective pressure in the cathode dark space is therefore increased and the maintaining-voltage decreased.

Can positive temperature coefficients be explained in a similar way by maintaining-voltage/pressure curves having positive gradients, or must additional factors be included in the theory? I have found, for example, that the temperature coefficient of the 85A2 tube at 5·5 mA, which is negative at low envelope temperatures, becomes positive at an envelope temperature of 130° C. The envelope temperature giving an effective pressure in the cathode dark space equal to the pressure at which the maintaining-voltage/pressure gradient changes from negative to positive is, however, much in excess of this.

Mr. G. C. Pope: I have been interested in the application of neon-argon-filled cold-cathode tubes as switching elements, in which higher argon contents are sometimes used. These give a larger difference between striking and maintaining voltage and hence a larger voltage can be developed into the load. With reference to the authors' comment that there is a general trend for low-argon-content tubes to have higher striking voltages, we have found with oxide-coated cathodes a minimum striking voltage in the region of 0·5–1% argon and the striking voltage rises on either side of this.

In connection with the temperature coefficient of maintaining voltage, Jacobs and Martin* showed that the exclusion of mercury vapour from argon tubes changed the temperature coefficient from negative to positive. Have the authors excluded mercury vapour from their experimental tubes, and, if not, to what extent does this contribute towards their results? Do the metastable mercury atoms add a further term to the impedance in the equivalent circuit?

Noise generation increases with argon content, as does also the current density. To what extent does the cathode surface uniformity contribute to noise generation?

Dr. R. O. Jenkins: I have been interested in the paper from the point of view of corona stabilizers, the geometry of which, quite fortuitously, happens to be very similar. One of the types has a 1 mm diameter anode and a 7 mm cathode, and the gas in this case is pure hydrogen at rather higher pressures than those used by the authors. The voltage drop is entirely confined to the neighbourhood of the anode, where the field is highest and the ionization is taking place. The currents are generally a fraction of a milliampere.

Such tubes always have a positive voltage/temperature coe cient. This ties up well with the fact that there is a posit variation in the voltage with pressure. On the other hand, th does not seem to be a direct correlation between the slope the voltage/pressure curve and the temperature coefficient. you regard the operating voltage as being $V_0(1 + aT)$, the va of a over a fairly wide range of pressure is fairly constant about 0·01% per deg C, so a 400-volt tube will have a change absolute magnitude of voltage for a given change in temperatu of only about a quarter of that for a 1 600-volt tube. The slop of the voltage/pressure curve at the two different points are n very different, that at 400 volts being, in fact, the greater.

Mr. G. F. Weston: I have been measuring the dischar characteristics of mixtures of helium and neon, and also helium and argon, with particular reference to the maintai ing potentials in geometries aimed at giving fairly unifor fields.

I find that in both systems of mixtures the voltage/curre curve depends on pressure, and, as a result, the plot of mai taining potential at a given current against the percentage neon or argon in helium is slightly different for different pressure This also applies to the minimum maintaining potentials. I ha also found that the geometry is very important, and if you use point anode and a plane cathode, instead of parallel planes, y again get a different curve of maintaining potential against g composition.

I would stress that one must specify the geometry and g pressure when dealing with gas mixtures, since the results m only apply to the particular conditions used. If the pressu and/or geometry are changed, a slightly different dependence gas composition may result.

Mr. R. E. Lake: The manufacturers' usual method measuring voltage jumps is by a current sweep technique, t jumps being displayed on an oscillograph.

One of the first essentials of a reference tube is, of cours stability during life, and I should be interested to know if D Benson intends to do, or has already done, some life tests these valves to examine the effect of argon content on th performance, or on their characteristics during life.

Mr. J. M. Glackin: Have the authors considered the effect elevated temperature on tube stability during life or shelf test

In one of the devices we manufacture we find that stabili on 100° C life test is slightly better than the stability at roo temperature. The only explanation which we can find for th is a change in argon content on room temperature life test due clean-up. At 100° C, however, providing the tube has be processed to withstand high temperatures, there is a change rate of clean-up due to evolution of argon occluded in t sputtered layer.

THE AUTHORS' REPLY TO THE ABOVE DISCUSSION

Dr. F. A. Benson and **Mr. P. M. Chalmers** (*in reply*): The main object of the present work was to find the influence of argon content on the characteristics of glow-discharge tubes, and high-stability reference tubes were used to reduce the spread in the results. The conclusion that 1% argon gives the most desirable performance was drawn after considering all the characteristics. We agree, however, that the noise is not then a minimum, but the variation of noise with argon content up to 1% is not large. Mr. Young points out that, at 1% argon, slight variations in argon content will affect the running voltage more than variations at 0·5%, and this is, of course, quite true. The spreads in the results at 1% and 0·5% argon, however, are not appreciably different and cannot be attributed to variations in the argon content alone. In Fig. 3, the 1% argon curve shows sensibly flat running voltage over the range 1–5 mA, while curv for lower argon contents have no flat regions.

In reply to Mr. Glackin, no thermal jumps were observ during the present investigations, but previous work usi coated-cerium cathodes indicated that jumps or steps in t voltage/temperature curves nearly always occurred in tub having similar steps in the running-voltage/current charact istics; these are thought to be due to sudden current-dens changes caused by non-uniform cathode surfaces.

Running-voltage/temperature characteristics were measur by immersing the tube in an oil bath of fairly large therm capacity, surrounded by a water jacket containing a heati element. Running voltage was recorded by a potentiometer

* *Journal of Applied Physics*, 1950, 21, p. 681.

ervals of 10° C over the range 20–90° C. At each tempera-
e, sufficient time was allowed for thermal equilibrium to be
ched and the running voltage to become constant. This
ied between 5 and 10 min and the temperature did not change
nificantly during the measuring period, when the heater
ment was switched off. We have recently investigated the
aracteristics of tubes containing neon plus 0·3% argon, and
d the influence of the geometry very small. We found that
noise was independent of the geometry, but increased with
s pressure at low tube currents. It is hoped to publish the
ults of this investigation shortly.

We agree with Mr. Weston that the shift between the minimum
nning-voltage curves of Fig. 2 is more likely to be due to a
fference in geometry than to a change in gas mixture caused
heavy sputtering. If the latter were the case, the difference
ould be more pronounced at low argon contents. The largest
screpancy occurs for argon contents exceeding 1% and this is
ssibly due to the start of an anode fall in these tubes.

Mr. Pope refers to the paper by Jacobs and Martin, who
liberately introduced a quantity of mercury into their tubes,
that a droplet remained after preparation. There was a
ssibility that mercury vapour was present in the tubes
amined, but subsequent investigations with additional tubes
the same type, free from contamination, produced character-
ics similar to those reported. We therefore concluded that, if
ercury vapour were present in the original tubes, it was
obably cleaned up in the sputtering process, or that the
nning effect due to the presence of mercury atoms was com-
etely masked by the argon admixture.

The impedance characteristics of a glow-discharge tube are more easily analysed when they are plotted in the form of impedance loci. The equivalent circuit can be identified from the form of the loci and can be regarded as a series combination of the static resistance (slope of static characteristic) and a number of parallel combinations of inductance and resistance. Each parallel combination has a time-constant associated with a particular process in the discharge mechanism. If mercury vapour is present in a 'pure-neon' tube, the lifetime of the neon metastable atoms will be reduced and the mercury ions formed by the Penning effect will contribute to the discharge current. Hence there will be an additional parallel LR combination in the equivalent circuit to account for this process. Similarly, if the mercury metastable atoms influence the discharge, another element will appear in the equivalent circuit, but it is difficult to see what role they would play in the discharge process.

We are interested to learn that Dr. Jenkins has observed positive voltage/temperature coefficients in corona stabilizers which could be traced back to the positive slope of the voltage/pressure curves. We are at present plotting voltage/pressure curves of glow-discharge tubes to see whether they have positive slopes at high argon contents.

Life characteristics of the tubes examined are also being studied, and we find that during the first 1 000 hours of operation the change of running voltage is small (0·1 volt), and independent of the argon content above 0·01%; during the first 5 000 hours the running voltage decreases by about 0·4 volt. These life tests are being performed at room temperature, the tubes being operated at a mean current. We have not so far carried out any stability tests at elevated temperatures.

DISCUSSION ON
'THE SQUARE-LOOP FERRITE CORE AS A CIRCUIT-ELEMENT'*

Mr. A. Kruithof (*communicated*): Dr. Lindsey does not take
e effect of the geometry of a ferrite core into account, although
is plays a major role at the field strengths usually encountered
, for instance, memory matrix operation.

Goodenough attributed the 'Gaussian' tail of the voltage
itput of a core to the random distribution of domains of reverse
agnetization.† At the field strengths which are usual in ferrite
re operation, however, the observed voltage waveform can be
rived from geometrical factors without necessarily assuming
random distribution.

Indeed, let us at first suppose that a perfectly regular distribu-
on of cylindrical domains of reverse magnetization nucleates,
r simplicity's sake, at one critical field strength H_n. This
ucleation phenomenom seems responsible for the initial
pike' of observed waveforms. In Permalloy cores another
echanism is responsible.

After nucleation the cylindrical domains have a radius r_0
ig. A) and expand at a speed which is proportional to the excess
the existing field strength over a fictitious field strength H_0
presenting the retarding forces. They expand at this speed
ntil $r = r_1$. Then the star-shaped form changes into another
linder with radius r_2, because the cylinder has a lower energy
ntent. This cylinder shrinks and disappears. An approximate
rve giving $\Delta\phi/\Delta t$ for this process is shown in Fig. B.

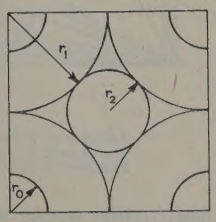


Fig. A.—Model for core reversal.

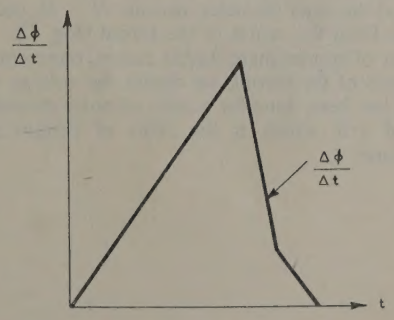


Fig. B.—Approximate elementary voltage output.

* Lindsey, C. H.: Monograph No. 327 M, February, 1959 (see 106 C, p. 117).
† Menyuk, N., and Goodenough, J. B.: 'Magnetic Materials for Digital Computer
mponents', *Journal of Applied Physics*, 1955, **26**, p. 8.

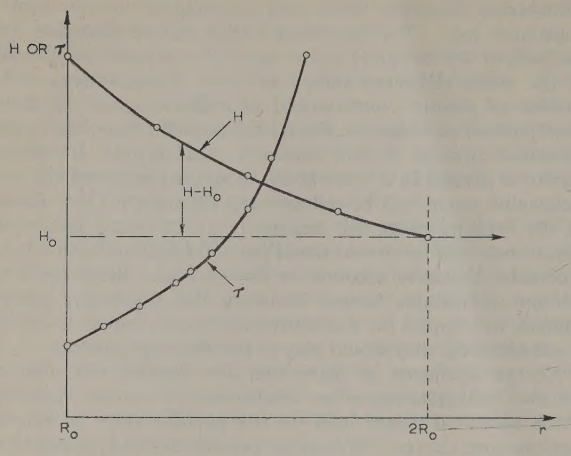


Fig. C.—Field strength H and reversal time τ as a function of the radius r when $H = 2H_0$ at $r = R_0$.

If no regular distribution of reverse domains is assumed, th approximate curve of $\Delta\phi/\Delta t$ of Fig. B has to be replaced by wav forms which could be obtained experimentally by observing th output of a core which had very approximately $R_1/R_0 \simeq 1$, c theoretically as Dr. Lindsey has done (theoretical waveform Fig. 7 in the monograph). By suitably adding, one again obtai the output for toroids with $R_1/R_0 \neq 1$. The result is that at th field strengths usually encountered and with, for exampl $R_1/R_0 = 1{\cdot}6$, a random distribution or a regular distributio makes little difference, as it is the geometry that dictates the shap of the voltage output.

Dr. C. H. Lindsey (*in reply*): By taking a rather idealized mod of the domain distribution, which leads to an output wavefor made up of three straight lines, and by then calling in the co geometry to round off the corners, Mr. Kruithof has, not su prisingly, produced a waveform with the same general appearanc as mine, which was based on a similar but not so highly idealize model.

Granted that I did not take the core geometry into accou (actually, the ratio of the outside to the inside diameter of th cores that I used was 1·5), this is only of importance when th

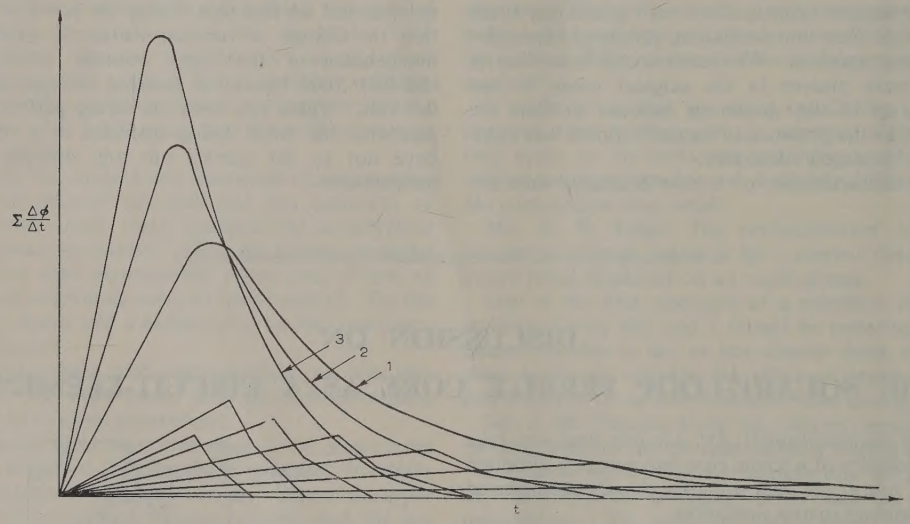


Fig. D.—Core voltage for a core with $R_1/R_0 = 1{\cdot}6$.

Addition of approximate $\Delta\Phi/\Delta t$ curves is shown for curve 1, for which $\bar{H} = 1{\cdot}8\ H_0$ at $r = R_0$.
For curve 2, $H = 2H_0$ at $r = R_0$.
For curve 3, $H = 2{\cdot}2H_0$ at $r = R_0$.

For a toroid the total voltage output is formed by the sum of all such processes going on throughout its cross-section. Near the inner diameter the process of expansion and shrinkage is faster than at the outer diameter, because $H \doteq H_0$ decreases with the distance from the centre of the toroid (Fig. C). By adding up a number of approximate $\Delta\phi/\Delta t$ curves, representative of the different parts of the toroid, we obtain the voltage output. In Fig. D this has been done for a ratio of outer diameter to inner diameter of 1·6, which is the ratio of certain well-known miniature cores.

actual field is not much greater than the pseudo-coercive force. Presumably Mr. Kruithof had this in mind when he referred to 'the field strengths which are usual in ferrite core operation'. Presumably he was referring to the field strengths used in coincident current matrix stores, but I was careful to point out, in my very first paragraph, that my theory did not pretend to cover this case. I think the evidence of Fig. 7 shows that the particular approximation I have made fits the facts very well over the relevant range, and I doubt very much whether Mr. Kruithof's model would fit the observed waveforms so well.

PROCEEDINGS OF THE INSTITUTION OF ELECTRICAL ENGINEERS

PART C—MONOGRAPHS, MARCH 1960

CONTENTS

Declaration on Fair Copying.—Within the terms of the Royal Society's Declaration on Fair Copying, to which The Institution subscribes, material may be copied from issues of the *Proceedings* (prior to 1949, the *Journal*) *which are out of print and from which reprints are not available.* The terms of the Declaration and particulars of a Photoprint Service afforded by the Science Museum Library, London, are published in the *Journal* from time to time.

Bibliographical References.—It is requested that bibliographical reference to an Institution paper should always include the serial number of the paper and the month and year of publication, which will be found at the top right-hand corner of the first page of the paper. This information should precede the reference to the Volume and Part.

Example.—SMITH, J.: 'Reflections from the Ionosphere', *Proceedings I.E.E.*, Paper No. 4001 R, December, 1954 (102 B, p. 1234).

★ There are many members and former members of The Institution who are finding life difficult. **When remitting your membership subscription,** please help them by sending a donation, or annual subscription preferably under deed of covenant, to

THE BENEVOLENT FUND

The object of the Fund is to afford assistance to necessitous members and former members (of any class) of The Institution of Electrical Engineers who have paid their subscriptions for at least five years consecutively or compounded therefor, and to the dependants of such members or former members.

Subscriptions and Donations may be sent by post to

THE INCORPORATED BENEVOLENT FUND OF THE INSTITUTION OF ELECTRICAL ENGINEERS, SAVOY PLACE, LONDON, W.C.2

or may be handed to one of the Local Hon. Treasurers of the Fund.

★ *Though your gift may be small, please do not hesitate to send it*

Local Hon. Treasurers of the Fund:

EAST MIDLAND CENTRE . . . . . *H. J. McLean*	SCOTTISH CENTRE . . . . *R. H. Dean, B.Sc.Tech.*
IRISH BRANCH . . . . . *A. Harkin, M.E.*	NORTH SCOTLAND SUB-CENTRE. . . . *P. Philip*
MERSEY AND NORTH WALES CENTRE *D. A. Picken*	SOUTH MIDLAND CENTRE . . . *H. M. Fricke*
TEES-SIDE SUB-CENTRE . . . . *W. K. Harrison*	RUGBY SUB-CENTRE . . . . . *P. G. Ross, B.Sc.*
NORTH-EASTERN CENTRE . . . . *J. F. Skipsey, B.Sc.*	SOUTHERN CENTRE . . . . . *J. E. Brunnen*
NORTH MIDLAND CENTRE *E. C. Walton, Ph.D., B.Eng.*	WESTERN CENTRE (BRISTOL) . . . *A. H. McQueen*
SHEFFIELD SUB-CENTRE . . . . . *F. Seddon*	WESTERN CENTRE (CARDIFF) . . . *E. W. S. Watt*
NORTH-WESTERN CENTRE . *E. G. Taylor, B.Sc.(Eng.)*	WEST WALES (SWANSEA) SUB-CENTRE *O. J. Mayo*
NORTH LANCASHIRE SUB-CENTRE . *H. Charnley*	SOUTH WESTERN SUB-CENTRE . . *W. E. Johnson*
NORTHERN IRELAND CENTRE . . *G. H. Moir, J.P.*	

Members are asked to bring to the notice of the Court of Governors any deserving cases of which they may have knowledge.

THE BENEVOLENT FUND

Published by The Institution, Savoy Place, London, W.C.2. *Telephone:* COVent Garden 1871. *Telegrams:* 'Voltampere, Phone, London.'
Printed by Unwin Brothers Limited, Woking and London.